# Functional genomics of monoterpenoid indole alkaloid biosynthesis in *Rauvolfia serpentina*

by

Paulo Emmanuel Cázares Flores

A thesis submitted to the Department of Biological Sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Faculty of Mathematics & Science, Brock University

St. Catharines, Ontario

## Abstract

Monoterpenoid indole alkaloids (MIAs) are a large and heterogeneous group of nitrogen-containing specialized metabolites produced by plants belonging to the Apocynaceae, Loganiaceae and Rubiaceae families. Many of these MIAs exhibit interesting biological activities and are currently used as pharmaceutical drugs to treat various medical conditions. Thus, the biosynthetic pathways responsible for their production have been extensively investigated. Recent advancements in large-scale DNA sequencing technologies have provided access to a vast collection of gene libraries. Here I have used a bioinformatics guided screen to identify candidate genes involved in MIA biosynthesis in *Rauvolfia serpentina* (L.) Benth. ex Kurz. Our large annotated transcriptome databases (www.phytometasyn.ca) were used as a source to mine genes. The identification of a *Catharanthus roseus* enzyme responsible for indole nitrogen methylation of an MIA intermediate of the vindoline pathway provided a query sequence to mine candidate genes responsible for *N*-methylation of other MIAs in *R. serpentina.* This led to the identification, molecular cloning and biochemical characterization of four enzymes catalyzing *N*-methylation. Two separate genes cloned from *R. serpentina* and *V. minor* both encoded enzymes displaying high affinity and specificity for picrinine, converting it to *N*-methyl-picrinine (ervincine) in the presence of *S*-adenosyl-L-methionine. The other two genes, cloned from *R. serpentina*, encoded enzymes involved in the final steps of ajmaline biosynthesis. Norajmaline *N*-methyltransferase catalyzed the indoline *N*-methylation of norajmaline to generate ajmaline, while ajmaline $N_{\beta}$-methyltransferase catalyzed the side chain *N*-methylation of ajmaline to generate $N_{\beta}$-methylajmaline, an unusual positively charged MIA found in Rauvolfia.

## Acknowledgements

I would especially like to thank my supervisor Vincenzo De Luca for giving me the opportunity to work in his lab and improve my skills as a scientist. I would also like to thank my committee members Charles Després and Douglas Bruce for their guidance and support throughout my research. I would also like to thank John T. Arnason and Jeffrey Atkinson for their great feedback to improve my documented research.

I would especially like to thank Dylan Levac and Fang Yu for their training and support during my first year as a graduate student, as well as the other past and current members of the De Luca lab: Brent Wiens, Vonny Salim, Antje Thamm, Kyung-Hee Kim, Alison Edge, Katie Woolfson, Michael Easson, Vince Qu and Zerihun Demissie.

I would like to thank all of the friends I have made along the way especially in the chemistry and biology departments: Mikel Ghelfi, Jordan Froese, Max Merilovich, Joao Fonseca, Yue Wu, Jan Bosák, Nina Bosáková, Min Pyo, Lucas Maddalena and Parthajit Mukherjee.

Finally, I would like to thank my family, and special thanks to Jacqueline Séguin for her unconditional love and support throughout all these years.

# Table of Contents

# List of Figures

## Chapter 3



**Supporting information**



## Chapter 4



**Supporting information**

## Appendix A

# List of Tables

## Chapter 3

**Supporting information**



## Chapter 4



**Supporting information**



## Appendix A

# List of Abbreviations

| | |
|---|---|
| 60S | 60S-ribosomal RNA subunit |
| AAE | Acetylajmalan esterase |
| BLAST | Basic local alignment search tool |
| cDNA | Complementary DNA |
| $C_T$ | Cycle threshold |
| *C. roseus* | *Catharanthus roseus* |
| CYP | Cytochrome P450 |
| DHVR | 1,2-dihydrovomilenine reductase |
| DNA | Deoxyribonucleic acid |
| DMAPP | Dimethylallyl diphosphate |
| *E. coli* | *Escherichia coli* |
| EDTA | Ethylenediaminetetraacetic acid |
| EST | Expressed sequence tags |
| E4P | Erythrose 4-phosphate |
| FPPS | Farnesyl diphosphate synthase |
| GES | Geraniol synthase |
| GPPS | Geranyl diphosphate synthase |
| G3P | Glyceraldehyde 3-phosphate |
| HPLC | High performance liquid chromatography |
| IPP | Isopentenyl diphosphate |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |

| | |
|---|---|
| IO | Iridodial oxidase |
| IR | Infrared |
| IS | Iridoid synthase |
| LAMT | Loganic acid *O*-methyltransferase |
| *mat*k | Maturase K |
| MEGA | Molecular evolutionary genetics analysis |
| MEP | 2-*C*-methyl-*d*-erythritol 4-phosphate pathway |
| MIA | Monoterpenoid indole alkaloid |
| MS | Mass spectrometry |
| MuSCLE | Multiple sequence comparison by log-expectation |
| MVA | Mevalonate pathway |
| NADPH | Nicotinamide adenine dinucleotide phosphate |
| NMT | *N*-methyltransferase |
| NNMT | Norajmaline *N*-methyltransferase |
| NMR | Nuclear magnetic resonance |
| OMT | *O*-Methyltransferase |
| ORF | Open reading frame |
| PAGE | Polyacrylamide gel electrophoresis |
| PCR | Polymerase chain reaction |
| PEP | Phosphoenolpyruvate |
| PNAE | Polyneuridine aldehyde esterase |
| PPP | Pentose phosphate pathway |
| RACE | Rapid amplification of cDNA ends |

| | |
|---|---|
| RNA | Ribonucleic acid |
| Rpm | Rotations per minute |
| RT-qPCR | Reverse-transcription quantitative PCR |
| Rf | Retention factor |
| *R. serpentina* | *Rauvolfia serpentina* |
| SABATH | Salicylic Acid, Benzoic Acid, Theobromine synthase family |
| SBE | Sarpagan bridge enzyme |
| SDS | Sodium dodecyl sulfate |
| SGD | Strictosidine β-glucosidase |
| SLS | Secologanin synthase |
| STR | Strictosidine synthase |
| TDC | Tryptophan decarboxylase |
| TLC | Thin-layer chromatography |
| γ-TLMT | γ-Tocopherol-like *N*- methyltransferase |
| TMT | γ-Tocopherol *C*- methyltransferase |
| UGT | 7-Deoxyloganetic acid glucosyltransferase |
| UPLC-MS | Ultra-performance liquid chromatography-mass spectrometry |
| UTR | Untranslated region |
| UV | Ultraviolet |
| *V. minor* | *Vinca minor* |
| VIGS | Virus induced gene silencing |

| | |
|---|---|
| VR | Vomilenine reductase |
| VS | Vinorine synthase |

# Chapter 1
# Introduction

*Rauvolfia serpentina* (L.) Benth. ex Kurz (Apocynaceae) has been extensively investigated over the last four decades due to its ability to produce biologically active monoterpenoid indole alkaloids (MIAs) such as ajmaline and reserpine, low molecular weight molecules used in the treatment of heart arrhythmias (Brugada et al., 2003) and high blood pressure (Anchor et al., 1955), respectively. The functional identification of the complementary metabolic pathway genes involved in ajmaline biosynthesis is very close to being complete (Wu et al., 2016). Here we have used a bioinformatics guided screen in order to select and functionally study candidate genes involved in ajmaline biosynthesis and other MIA pathways.

**Chapter 2** provides a literature review of plant specialized metabolism focusing on the elucidation and functional characterization of MIA pathway genes from the Apocynaceae plant family and how scientists have used gene libraries that have recently become available through large-scale DNA sequencing projects for performing functional genomics and MIA pathway elucidation in different plant species.

**Chapter 3** is a manuscript recently published in *Plant Physiology*. Here I have documented part of the discovery of a new family of enzymes responsible for the *N*-methylation of MIA from the Apocynaceae plant family, the γ-tocopherol-like *N*-methyltransferase (γ-TLMT) family. Moreover, we provided an interesting approach for the identification of enzyme substrates based on targeted enzyme-activity profile.

**Chapter 4** is a manuscript that was recently accepted for publication in *The Plant Journal* and will appear in the next issue. Here I have documented that enzymes

responsible for the final steps in the biosynthesis of ajmaline belong to the γ-TLMT family. Norajmaline *N*-methyltransferase carries out the indoline *N*-methylation of norajmaline generating ajmaline, and ajmaline $N_{\beta}$-methyltransferase carries out the side chain *N*-methylation of ajmaline producing a quaternary ammonium MIA, $N_{\beta}$-methylajmaline. We successfully identified the regiospecific methylation of each recombinant γ-TLMT by analyzing their reaction products through tandem mass spectrometry (MS-MS). Additionally, we provided an approach for gaining access to substrates through microbial biotransformation.

**Chapter 5** provides a general discussion of the research documented in this thesis.

**Chapter 6** provides the main conclusions of this study and future prospects for characterizing the remaining pathway genes involved in ajmaline biosynthesis.

### 1.1 Co-Authorship

Described below are the contributions from each author to multi-authored chapters.

**Chapter 3**: A picrinine *N*-methyltransferase belongs to a new family of γ-tocopherol-like *N*-methyltransferases (γ-TLMTs) found in medicinal plants that make biologically active monoterpenoid indole alkaloids.

This work was initiated by Dylan Levac and Vincenzo De Luca and has been published in Plant Physiology (*Plant Physiology*, 2016, ***170***, 1935–1944). The manuscript was written by Vincenzo De Luca, Dylan Levac and Paulo Cázares Flores. Vincenzo De Luca, Dylan Levac and Paulo Cázares Flores were responsible for experimental design. Dylan Levac, Paulo Cázares Flores and Fang Yu performed the experiments. Dylan

Levac and Fang Yu performed the initial bioinformatics-guided searches for candidate genes and their molecular cloning, functional expression of enzymes and substrate elucidation based on targeted enzyme-activity profiles and substrate specificity assays. Paulo Cázares Flores performed the expression and purification of recombinant enzymes, determination of enzyme kinetic parameters, substrate specificity assays, and structure analysis of enzyme reaction products by ultraviolet and infrared spectroscopy and mass spectrometry.

**Chapter 4**: *Rauvolfia serpentina N*-methyltransferases involved in ajmaline and $N_{\beta}$-methylajmaline biosynthesis belong to a gene family derived from γ-tocopherol *C*-methyltransferase.

Work from this chapter has been accepted for publication in The Plant Journal (*The Plant Journal* April 2016, doi: 10.1111/tpj.13186). The manuscript was written by Paulo Cázares Flores and Vincenzo De Luca. Vincenzo De Luca, Dylan Levac and Paulo Cázares Flores were responsible for experimental design. Paulo Cázares Flores and Dylan Levac performed the experiments. Dylan Levac performed initial the bioinformatics-guided search for candidate genes. Dylan Levac and Paulo Cázares Flores performed the molecular cloning of candidate genes. Paulo Cázares Flores performed the expression and purification of recombinant enzymes, the determination of enzyme kinetic parameters and substrate specificity, the structural analysis of enzyme reaction products by ultraviolet spectroscopy and tandem mass spectrometry (MS/MS), and the enzyme activity, relative transcript abundance, and phytochemical analyses of organs and leaf developmental stages.

## 1.2 Research goals and objectives

Over the last decade, the advent of large-scale DNA sequencing technologies has facilitated access to a vast number of gene libraries allowing the use of bioinformatics tools to elucidate plant specialized metabolic pathways (Salim and De Luca, 2013). The goal of this thesis was to use bioinformatics guided screening in order to identify and study candidate metabolic pathway genes from our large annotated transcriptomic databases of medicinal plants (http://www.phytometasyn.com/; Facchini et al., 2012; Xiao et al., 2013) with special emphasis in *Rauvolfia serpentina* genes related to the ajmaline biosynthetic pathway. Recently, a new family of *S*-adenosyl-L-methionine (AdoMet)-dependent enzymes phylogenetically related to the γ-tocopherol *C*-methyltransferase family has been identified in several members of the Apocynaceae plant family (Liscombe et al., 2010; Levac, 2013). This novel gene family, the γ-TLMT, appears to be responsible for the hundreds of indole and side chain *N*-methylated MIAs within Apocynaceae (Levac, 2013). The objective of this thesis was the elucidation and molecular and biochemical characterization of the remaining unknown metabolic pathway steps involved in ajmaline biosynthesis. Furthermore, we focused on the two terminal reactions involving AdoMet-dependent enzymes: norajmaline and ajmaline *N*-methyltransferases, and determined whether these two enzymes belong to the γ-TLMT family.

# Chapter 2
# Literature review

## 2.1 Plant secondary/specialized metabolites

All higher plants are known to biosynthesize low molecular weight organic compounds that do not appear to play an essential role in growth and development, yet their presence is thought to provide defenses against insects, fungi and other pests (Theis and Lerdau, 2003). These compounds have been categorized as secondary metabolites, natural products or specialized metabolites that serve as defensive chemicals against herbivory or microbial infection (Croteau et al., 2000; Hartmann, 2004), as attractants for pollinators (Pichersky and Gershenzon, 2002), as UV protectants (Dixon and Paiva, 1995), or as allelochemicals that influence competition against surrounded plant species (Liu and Lovett, 1993).

Interestingly, many plant secondary metabolites have been exploited by human societies and currently some of these compounds are being used as dyes, fibers, oils, flavoring agents, perfumes, and pharmaceutical drugs (Hartmann, 2007). These commercial applications have fueled intense research to determine the biochemical pathways responsible for their synthesis, their allocation within the plants, and investigate potential targets for genetic engineering in order to increase the accumulation of the desired secondary metabolites (Verpoorte and Alfermann, 2000).

### 2.1.1 Discovery of secondary metabolism pathways

Although in the late 1800's plant secondary metabolites were thought to be merely metabolic waste products, their remarkably potent biological activities and complex chemical structures prompted rapid recognition of their importance in producing

organisms by chemists and biologists (Herbert, 1989). Currently more than 150,000 heterogeneous chemical structures of plant natural products have been described and elucidated on an ongoing process of discovery (Harvey et al., 2015; Croteau et al., 2000; De Luca and St. Pierre, 2000). In addition, biologically active natural products stimulated the development of powerful separation and spectroscopic methods such as high performance liquid chromatography (HPLC), mass spectroscopy (MS) and nuclear magnetic resonance (NMR) spectroscopy commonly used for detection, structure elucidation and purification of various secondary metabolites.

Initial mechanistic studies of plant natural product biosynthesis were focused on determining their biosynthetic origins by tracking the incorporation patterns of radioactive isotopes ($^{14}C$ and $^{3}H$) from labelled-precursor/intermediates in feeding experiments (Herbert, 1989). These experiments led to the conclusion that secondary metabolites were synthesized *de novo* from simple compounds produced in primary metabolism through an elaborate and organized series of reactions catalyzed by specialized enzymes (Herbert, 1989). These specialized enzymes furnished the core structure of primary starting materials by adding new functional groups, or by removing already present ones. Other specialized enzymes catalyzed coupling reactions or the formation of new cyclic structures. It is important to mention that primary metabolism comprises biochemical pathways involved in the elaboration of metabolites that support growth, development and reproduction. Furthermore, while primary metabolic pathways tend to be common to all living organisms (Cavalier-Smith, 1992), specialized metabolic pathways are usually restricted to specific taxonomic families (Wink, 2003; Pichersky and Gang**,** 2000).

Plant secondary metabolites can be classified according to their biosynthetic origins and can be divided into three major classes: terpenoid, alkaloid and phenolic compounds. The terpenoids constitute the most structurally diverse class of secondary metabolites with over 34,000 known chemical structures (Croteau et al., 2000). Terpenoid molecules are derived from the five-carbon intermediates isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) which in plants can be synthesized via a cytosolic mevalonate (MVA) pathway or a plastidic 2-*C*-methyl-*d*-erythritol 4-phosphate (MEP) pathway (Vranová et al., 2013). Alkaloids are a highly diverse group of nitrogen-containing heterocyclic molecules with approximately 12,000 described structures (Buckingham et al., 2010). Alkaloids are mainly biosynthesized from the amino acids, phenylalanine, tyrosine, tryptophan, lysine and ornithine (De Luca and St. Pierre, 2000; Facchini, 2001). Finally, phenolics are derived from the common intermediate 4-coumaroyl Coenzyme A that is derived from phenylalanine or tyrosine units. Phenolics comprise a wide array of over 2,500 identified compounds (Croteau et al., 2000; Vogt, 2010).

The use of isotope labelling to track the steps involved in a biosynthetic pathway has also been important for the identification of possible substrates for the individual biochemical reactions involved and for developing enzyme assays that corroborate the proposed pathway (Herbert, 1989). Enzyme assays were then used for purification and biochemical characterization of enzymes involved in specialized metabolism from plant sources since the purification could be tracked through column chromatographic steps including hydrophobic interaction, size exclusion, ion exchange and affinity chromatography (Herbert, 1989; Hartmann, 2007). After purification to homogeneity, the

target enzymes could be submitted to partial proteolysis, amino acid sequencing of peptides and the peptide sequence information was used to design oligonucleotide primers for screening cDNA libraries and cloning sequences encoding their respective open reading frames (ORF). Further developments in recombinant DNA technology allowed the heterologous expression of clones that produced relevant enzymes in microorganism such as bacteria and yeast. This facilitated purification of recombinant proteins that could be submitted to detailed biochemical and structural analysis. More recently, the advent of large-scale DNA sequencing projects marked a new era in the biochemical and molecular characterization of enzymes involved in secondary metabolism by facilitating the selection of candidate genes through comparative bioinformatics (De Luca et al., 2012). Such tools for candidate gene identification combined with plant gene silencing techniques involving RNA interference (Baulcombe, 1999; Meister and Tuschl, 2004) have assisted in the identification of genes involved in specialized metabolism by providing a reverse-genetics approach for their characterization (Liscombe and O'Connor, 2011; Salim and De Luca, 2013).

After decades of intense research, many of these secondary metabolic pathways have been characterized at the molecular and biochemical level, providing a large collection of cDNAs encoding the biosynthetic enzymes and regulatory proteins required for their formation. Pioneer model systems for studying monoterpenoid indole alkaloid biosynthesis include those of *Catharanthus roseus* (L.) G. Don and *Rauvolfia serpentina* (De Luca et al., 2012), while the *Papaver somniferum* (L.) model system has been used to study the biosynthesis of benzylisoquinoline alkaloids (Facchini, 2001). The following

sections focus on the discovery and functional analysis of enzymes involved in monoterpenoid indole alkaloid biosynthesis.

## 2.2 Monoterpenoid indole alkaloids

MIAs are a large and diverse group of nitrogen-containing heterocyclic secondary metabolites with over 2,000 that have been structurally characterized (Buckingham et al., 2010). These molecules have been extensively investigated due to their powerful biological activities and their pharmaceutically valuable applications. Some MIAs are used as chemotherapeutic agents in different cancer treatments (Noble, 1990), as hypotensive compounds in the treatment of high-blood pressure (Achor et al., 1955), and in the diagnosis and therapy of several heart arrhythmias (Brugada et al., 2003). The biochemical pathways responsible for their biosynthesis are mostly restricted to plants belonging to the Apocynaceae, Loganiaceae, and Rubiaceae families (Facchini and De Luca, 2008; Kisakurek et al., 1983).

Most MIAs are derived from the common intermediate strictosidine (Nagakura et al., 1979) which is produced by the condensation of tryptamine derived from the decarboxylation of the amino acid tryptophan (De Luca et al., 1989) with secologanin, a cyclic-monoterpene glycoside. This condensation is catalyzed by the enzyme strictosidine synthase (STR) (Treimer and Zenk, 1979) through a stereoselective Pictet-Spengler reaction mechanism in order to yield a β-carboline product (Maresh et al., 2008). Strictosidine β-glucosidase (SGD) (Luijendijk et al., 1998) cleaves the glucose moiety of strictosidine to produce an unstable aglycone molecule which spontaneously leads to the formation of a series of reactive intermediates that serve as starting materials for the biosynthesis of different MIA backbones/groups including the sarpagan, corynanthe,

aspidosperma and iboga types (**Figure 2-1**) (O'Connor and Maresh, 2006). The enzymes catalyzing these complex rearrangements of the terpenoid moiety to produce the different MIA backbones are restricted to specific genera within the above mentioned plant families (Salim and De Luca, 2013); this feature highlights the phylogenetic origin and evolution of specific secondary metabolic pathways (Wink, 2003). Interestingly, *Catharanthus roseus*, a member of the Apocynaceae family, is the only known plant species to possess the biosynthetic enzymes involved in the formation of the four mentioned MIA backbones (O'Connor and Maresh, 2006), thus making *C. roseus* a prominent system to study MIA biochemistry (Salim and De Luca, 2013).

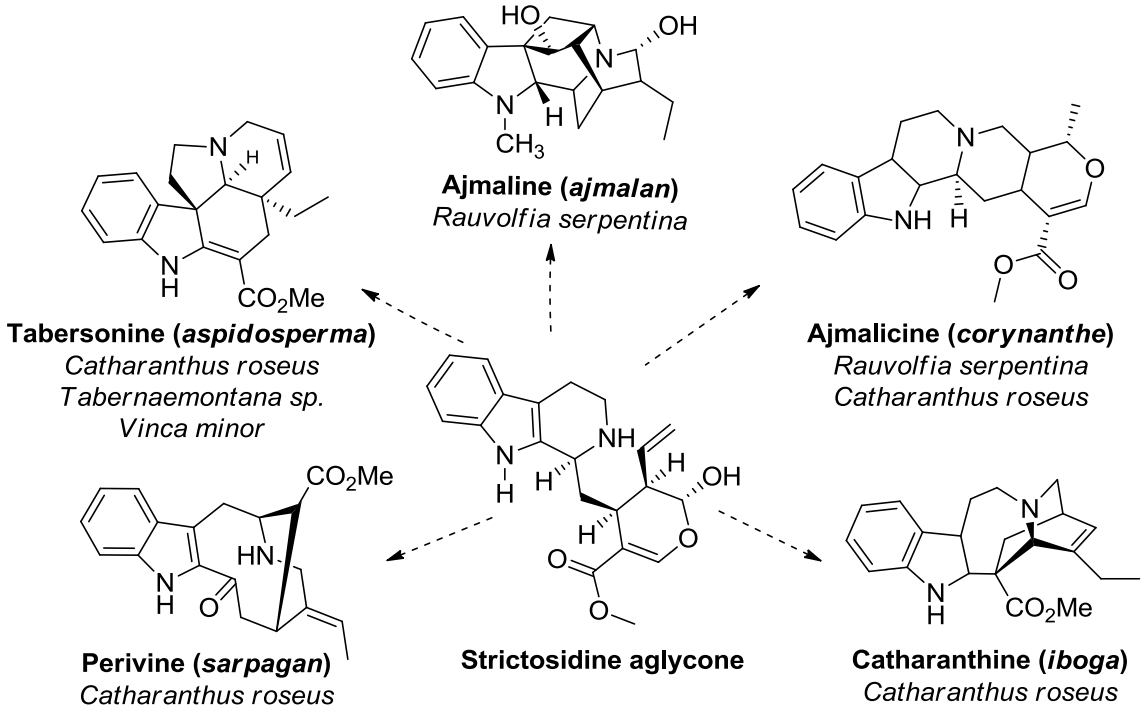



**Figure 2-1: MIAs found in *Catharanthus roseus* and other members of the Apocynaceae plant family**. Different monoterpenoid indole alkaloid backbones are produced from the common strictosidine aglycone intermediate. Dashed arrows indicate multiple enzymatic reactions.

### 2.2.1 Early pathway – biosynthesis of strictosidine

Secologanin (**Figure 2-2**), a cyclic-monoterpene glycoside, also classified as a secoiridoid, is an important building-block in the biosynthesis of MIAs, quinoline alkaloids and pyrroloquinoline alkaloids (Tietze, 1983). The initial biogenetic precursor of secologanin, isopentenyl-diphosphate (IPP), is produced mainly via the MEP pathway; analysis of secologanin $^{13}$C-NMR data obtained from feeding experiments of *C. roseus* cell suspension cultures supplemented with [1-$^{13}$C]-glucose showed isotopic enrichment of carbons typical of the plastid localized MEP pathway instead of the cytosolic mevalonate pathway (Contin et al., 1998) in the assembly of the $C_{10}$ monoterpene-moiety. Furthermore, geranyl-diphosphate synthase (GPPS) which catalyzes the condensation of dimethylallyl diphosphate (DMAPP) and IPP to assemble the $C_{10}$ monoterpene backbone has been immunolocalized in plastids from *A. thaliana* (Bouvier et al., 2000) and *Catharanthus roseus* leaves (Thabet et al., 2011). These studies highlight the importance of this organelle in monoterpene biosynthesis in *A. thaliana*, *Catharanthus roseus* and in general in other plants.

The next step in the pathway (**Figure 2-2**) towards the biosynthesis of secologanin involves the removal of the pyrophosphate group from geranyl diphosphate by the enzyme geraniol synthase (GES) (Iijima et al., 2004). The enzyme GES was initially purified from sweet basil glands in a four step purification protocol (Iijima et al., 2004). Moreover, GES was shown to have a similar catalytic mechanism to other terpene synthases where the reaction occurs through the addition of a hydroxyl group to a carbocation intermediate instead of hydrolyzing the phosphoester bond (Iijima et al., 2004). Geraniol is then further hydroxylated at the C-10 position by the action of geraniol

10-hydroxylase (G10H), an enzyme belonging to the cytochrome P450 (CYP) superfamily (Collu et al., 2001). G10H represents one of the few examples where an endoplasmic reticulum bound CYP enzyme was successfully purified to apparent homogeneity to enable partial amino acid sequencing for further cloning and recombinant expression of its cDNA (Collu et al., 2001). G10H was purified from cell suspension cultures of *C. roseus* where its gene expression was up regulated by the addition of the plant hormone methyl jasmonate to the culture (Collu et al., 2001).

This acyclic monoterpene alcohol (10-hydroxygeraniol) is then further oxidized to 10-oxogeranial, which is the biogenetic precursor of all iridoids, by 10-hydroxygeraniol oxidoreductase (10HGO) (**Figure 2-2**). This enzyme was purified to homogeneity after extraction from *R. serpentina* cell suspension cultures (Ikeda et al., 1991). The substrate is converted by 10HGO to 10-oxogeranial, a dialdehyde monoterpene substrate for the biosynthesis of the cyclic monoterpene iridodial. Iridoid synthase (IS) converts 10-oxogeranial to a mixture of cyclic monoterpenes that include *cis-trans*-nepetalactol in equilibrium with *cis-trans*-iridodial, through a novel reductive cyclization reaction mechanism (Geu-Flores et al., 2012). The *IS* gene was identified using the transcriptomic data based on its similar expression pattern to other genes involved in secologanin biosynthesis in *C. roseus* and combined with virus induced gene silencing and targeted metabolomics analysis (Geu-Flores et al., 2012; Góngora-Castillo et al., 2012). Amino acid sequence analysis of IS revealed high amino acid identity to progesterone-5β-reductase, a gene that belongs to the short-chain reductase gene family. Moreover, IS required the cofactor NADPH for enzyme activity, reflecting its assignment to this gene family (Geu-Flores et al., 2012). The following enzymatic step involves the conversion of

nepetalactol/iridodial to the bicyclic carboxylated monoterpene 7-deoxyloganetic acid by the action of iridodial oxidase (IO), a CYP enzyme which carries out a 3 step oxidation reaction (Salim et al., 2014). The IO candidate gene was discovered through a bioinformatics approach involving comparative transcriptome analysis of CYPs from several plant species known to biosynthesize MIAs and the secologanin-producing plant *Lonicera japonica* (Caprifoliaceae family) (phytometasyn.com, Facchini et al., 2012; Xiao et al., 2013). Candidate CYP genes were further screened against *L. japonica* that produces secologanin but not MIAs in order to narrow down candidate genes involved in secologanin biosynthesis (Salim et al., 2014). Virus induced gene silencing (VIGS) of *IO* transcripts confirmed its involvement in 7-deoxyloganetic acid synthesis and further recombinant expression of *IO* cDNA in yeast supported its biochemical function (Salim et al., 2014). The free hydroxyl group of 7-deoxyloganetic acid is then glycosylated by 7-deoxyloganetic acid glucosyltransferase (UGT) to produce 7-deoxyloganic acid (Asada et al., 2013). This enzyme utilizes uridine diphosphoglucose as a cofactor to transfer the glucose moiety into the iridoid substrate. The full-length cDNA of *UGT* was cloned based on oligonucleotide primers designed against a conserved amino acid sequence motif present in members of the plant secondary product glycosyltransferase family located near the C-terminus of the protein by using rapid amplification of cDNA ends (RACE) (Asada et al., 2013). The substrate specificity of this UGT enzyme revealed the appropriate biochemical order of the reactions to produce secologanin (Asada et al., 2013) depicted in **Figure 2-2**.

The secoiridoid 7-deoxyloganic acid is then further modified by the addition of a hydroxyl group at the C-7 position, the enzyme responsible for this reaction, 7-

deoxyloganic acid 7-hydroxylase (DL7H), belongs to the CYP family and was functionally identified based on a similar bioinformatics approaches as those described for the identification of the IO gene (Salim et al., 2013; Salim et al., 2014). VIGS of *DL7H* transcripts in *C. roseus* seedlings directly correlated with the accumulation of 7-deoxyloganic acid substrate and decreased levels of secologanin and MIAs compared to control plants (Salim et al., 2013). Furthermore, recombinant expression of *DL7H* cDNA in *Saccharomyces cerevisiae* microsomes confirmed its biochemical function and strict substrate specificity for 7-deoxyloganic acid in order to produce loganic acid (Salim et al., 2013). The carboxylic acid group of loganic acid is then methylated by loganic acid *O*-methyltransferase (LAMT) to generate loganin. The *LAMT* gene was cloned (Murata et al., 2008) based on its amino acid sequence similarity to members of the *S*-adenosyl-L-methionine-dependent carboxyl methyltransferases from the SABATH (Salicylic Acid, Benzoic Acid, Theobromine synthase) family (D'Auria et al., 2003). Recombinant expression of *LAMT* cDNA in *Escherichia coli* produced a protein with high specificity for loganic acid, whereas other iridoids such as 7-deoxyloganic and loganetic acids were not turned over (Murata et al., 2008). Additionally, LAMT transcripts and enzyme activity were enriched in the leaf epidermis of *C. roseus* compared to the whole leaf (Murata et al., 2008). The last step in the biosynthesis of secologanin involves the oxidative ring cleavages of loganin by a CYP enzyme to yield the final secoiridoid molecule (Irmler et al., 2000). This CYP enzyme, named secologanin synthase (SLS), was isolated from a cDNA library generated from *C. roseus* cell suspension cultures grown in a media designed to induce MIA biosynthesis (Vetter et al., 1992). Furthermore, *in situ* RNA hybridization studies localized *SLS* transcripts in leaf epidermal cells (Irmler

et al., 2000). These localization studies highlighted the importance of leaf epidermal cells as a site for completing the assembly of secologanin and for MIA biosynthesis.

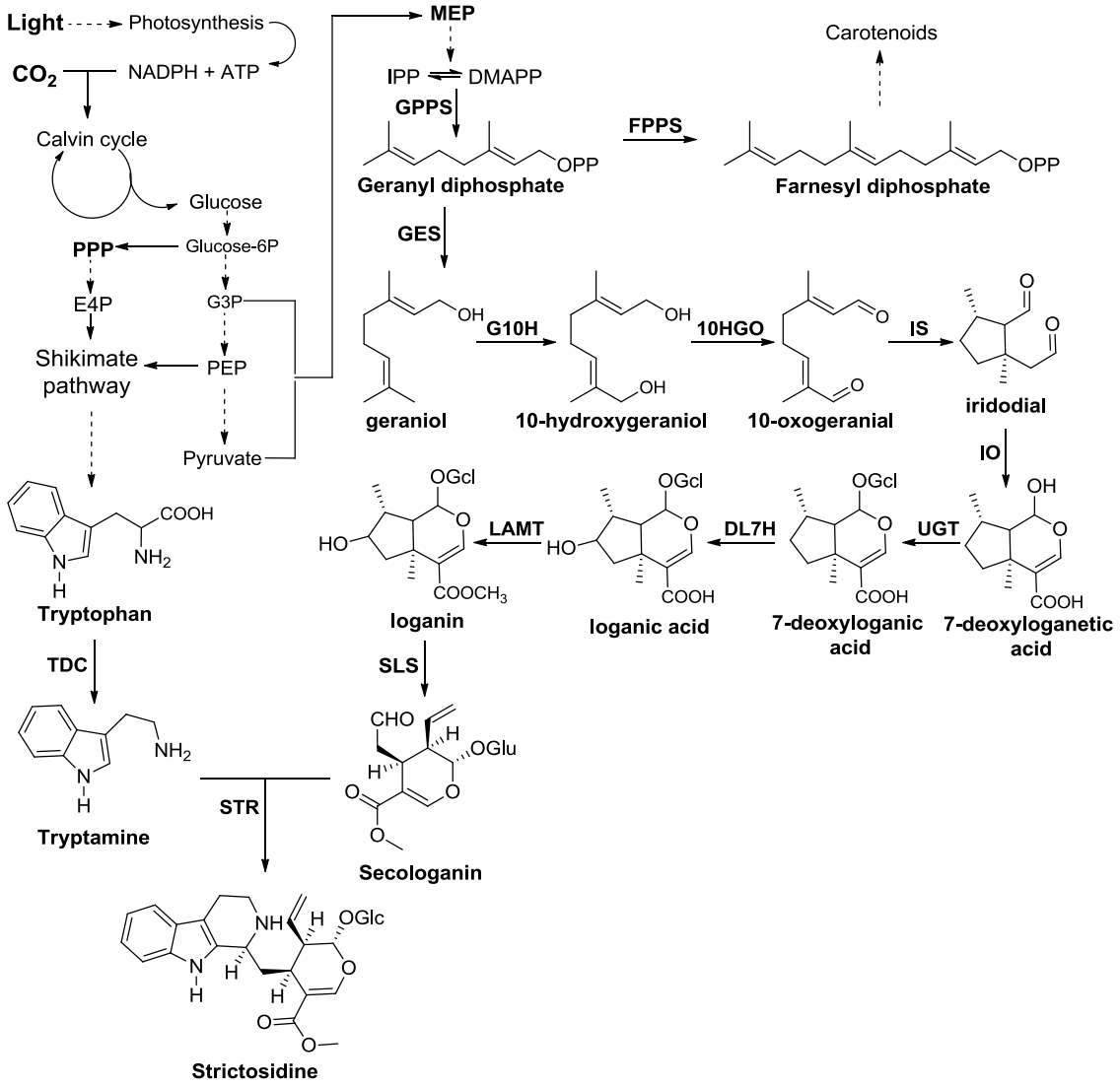



**Figure 2-2**: **Condensed full-scheme of strictosidine biosynthetic pathway.** Initial biogenetic precursors are produced in primary metabolic pathways: Calvin cycle, glycolysis, pentose phosphate pathway (PPP), shikimate pathway and 2-*C*-methyl-d-erythritol 4-phosphate (MEP) pathway. Abbreviations: glyceraldehyde 3-phosphate (G3P), phosphoenolpyruvate (PEP), erythrose 4-phosphate (E4P), geranyl diphosphate synthase (GPPS) and farnesyl diphosphate synthase (FPPS). Scheme reconstructed from: Salim and De Luca, 2013.

### 2.2.2 Late pathway– biosynthesis of ajmaline in *Rauvolfia serpentina*

*Rauvolfia serpentina* (Apocynaceae) has been extensively used as a platform to study the biosynthesis of sarpagan and ajmalan type MIAs (Stöckigt and Panjikar, 2007). Detailed chemical analyses of this and related species have led to the structural characterization of more than twenty-five MIAs (Klyushnichenko et al., 1995; Stöckigt et al., 1997), mostly belonging to the ajmalan group of MIAs.

Due to the extremely slow growth of differentiated intact *R. serpentina* plants, investigators in the early 1980's opted for a more amenable system for biochemical characterization of the specialized enzymes involved. The system involved the use of the well-established MIA accumulating, rapidly growing cell suspension cultures which provided an excellent source of enzymes for studying the biosynthesis of *R. serpentina* MIAs (Ruppert et al., 2005a). Phytochemical analyses of *R. serpentina* cell suspension cultures confirmed that they accumulate a number of MIAs found in the parent plants and established the usefulness of these *in vitro* cultures as MIA enzyme sources (Stöckigt et al., 1981). Thus all studies on the biochemical and molecular characterization of the ajmaline pathway, one of the major MIAs of *R. serpentina,* were performed with cell suspension cultures (Stöckigt et al., 1981).

The initial modification of the strictosidine aglycone to initiate ajmaline biosynthesis requires the generation of the sarpagan ring system between the indole C-5 and terpene C-16 leading to the formation of polyneuridine aldehyde, a reaction carried out by sarpagan bridge enzyme (SBE) (Schmidt and Stöckigt, 1995). While current experiments propose geissoschizine as the substrate of SBE, it is not clear whether this enzyme could use 4,21-dehydrogeissoschizine or perhaps a different iminium ion such as

4,5-dehydrogeissoschizine for catalysis (Ingham et al., 2012; O'Connor and Maresh, 2006). The activity of SBE has been detected in microsomal preparations from cell suspension cultures where the reaction was dependent on molecular oxygen and NADPH (Schmidt and Stöckigt, 1995). While these results suggested that SBE may be a member of the CYP superfamily, no *SBE* gene has been identified and functionally characterized. The next step in the pathway involves hydrolysis of the polyneuridine aldehyde methyl ester group by polyneuridine aldehyde esterase (PNAE) to form a reactive carboxylic acid that spontaneously decarboxylates to yield 16-epivellosime (Pfitzner and Stöckigt, 1983). PNAE represents the first committed step in the biogenesis of both sarpagan and ajmalan type MIA backbones. Spontaneous epimerization of 16-epivellosime to vellosime channels this latter molecule exclusively into the sarpagan type of MIAs (Pfitzner and Stöckigt, 1983). PNAE was purified to homogeneity and fragments of the purified protein were used for oligonucleotide primer design leading to the isolation of the PNAE cDNA from a *R. serpentina* library (Dogru et al., 2000). Additionally, structural analysis of the recombinant enzyme showed that PNAE belongs to the α/β hydrolase superfamily and requires a Ser, His, Asp catalytic triad for activity (Mattern-Dogru et al., 2002; Yang et al., 2009).

Formation of the sixth ring system between C-7 and C-17 generates the ajmalan type MIA backbone (Pfitzner et al., 1986). This reaction is catalyzed by vinorine synthase (VS), a member of the BAHD acyl-CoA-dependent acyltransferase superfamily. This enzyme uses 16-epivellosime (*endo* configuration of the aldehyde group) exclusively as substrate in order to generate the *O*-acetylated indolenine compound vinorine from Acetyl CoA (Ma et al., 2005). The molecular cloning and functional characterization of

VS involved purification of the enzyme to homogeneity from *R. serpentina* cell suspension cultures and protein sequencing to obtain appropriate oligonucleotide primers for isolation of the relevant gene from a cDNA library (Bayer et al., 2004). Further modification of vinorine involves the hydroxylation of C-21 by a microsomal CYP enzyme (Falkenhagen and Stöckigt, 1995) to produce vomilenine. Vomilenine reductase (VR) and 1,2-dihydrovomilenine reductase (1,2-DHVR) are NADPH-dependent reductases that catalyze the saturation of the 1,2-indolenine double bond (von Schumann et al., 2002), and 19,20 double bond convert it to yield 1,2-dihydrovomilenine and 17-*O*-acetylnorajmaline (1,2,19,20-tetrahydrovomilenine) (Gao et al., 2002), respectively. These two reductases have been purified to homogeneity and partial protein sequences were obtained for both enzymes (Ruppert et al., 2005a). Nevertheless, their putative cDNA clones still remain to be tested for activity, e.g. recombinant expression in heterologous systems such as in *Escherichia coli* or *S. cerevisiae*.

The next step along the ajmaline biosynthetic pathway involves hydrolysis of the 17-*O*-acetylated norajmaline by the enzyme acetyl ajmalan esterase (AAE). Substrate specificity assays for AAE demonstrated a strict requirement of the saturated 19-20 bond produced by 1,2DHVR, revealing the specific order of the biochemical reactions (Polz et al., 1987). The cDNA of AAE was functionally expressed in *Nicotiana benthamiana* leaves using a virus-based plant protein expression system and enzyme assays using leaf extracts confirmed its biochemical activity (Ruppert et al., 2005b). The last step involves an *S*-adenosyl-*L*-methionine-dependent methyltransferase (Stöckigt et al., 1983) which catalyzes the indoline nitrogen methylation of norajmaline to produce ajmaline. Although the biochemical activity of this *N*-methyltransferase has been detected in cell-free

extracts from *R. serpentina* cell suspension cultures, the enzyme has not been further characterized at the biochemical and molecular level.

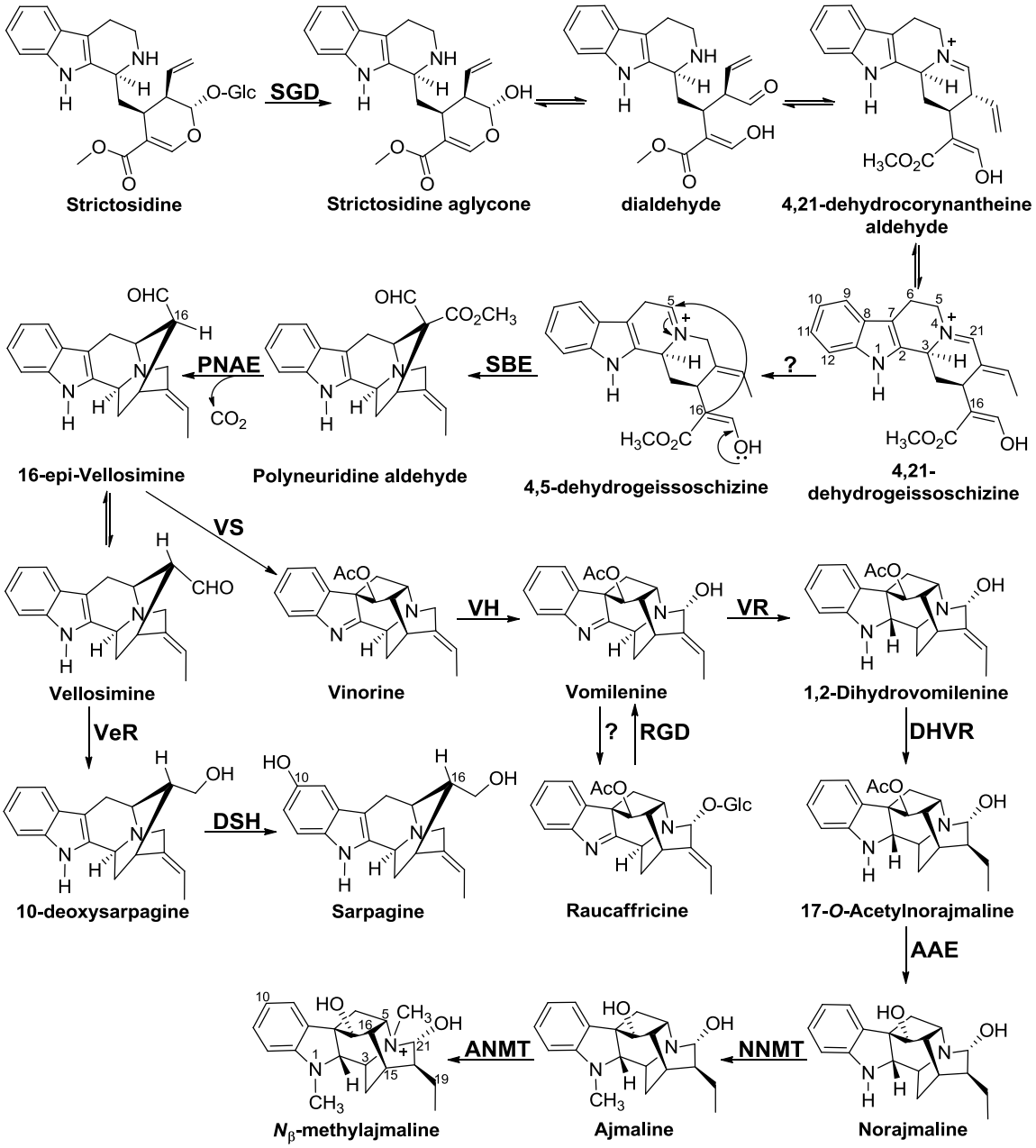



**Figure 2-3: Biosynthesis of sarpagan and ajmalan MIA backbones in *R. serpentina*.** Initial biogenetic intermediated 4,21-dehydrogeissoschizine is produced through the cleavage of the glucose moiety from strictosidine, reaction carry out by strictosidine β-D-glucosidase (SGD). Unknown enzyme generates potential substrate (4,5-

dehydrogeissoschizine) for the sarpagan bridge enzyme (SBE), an unidentified member of the cytochrome P450 (CYP) family. Polyneuridine aldehyde esterase (PNAE) produces the important branching intermediate, 16-epi-vellosimine, between sarpagan and ajmalan backbones. Vinorine synthase (VS) mediates the formation of the ajmalan-type backbone. Vomilenine is produced through the C-21 hydroxylation of vinorine by an unknown CYP: vinorine hydroxylase (VS). Reduction of the indolenine 1,2 and terpene 19,20 double bonds from vomilenine are carry out by two NADPH-dependent reductases, vomilenine reductase (VR) and 1,2-dihydrovomilenine reductase (DHVR). Acetylajmalan esterase (AAE) generates norajmaline that is indoline *N*-methylated by norajmaline *N*-methyltransferase (NNMT). The side chain nitrogen of ajmaline is methylated by ajmaline *N*-methyltransferase (ANMT) producing a quaternary ammonium MIA. Vellosimine reductase (VeR) and 10-deoxysarpagine hydroxylase (DSH) contribute to the biosynthesis of sarpagine. Metabolic pathway scheme was reconstructed from: O'Connor and Maresh, 2006, and Stöckigt and Panjikar, 2007.

In summary, most of the ajmaline biosynthetic pathway has been well established at the molecular and biochemical level. Cell suspension cultures of *R. serpentina* provided the optimum system to characterize individual enzymatic reactions leading to the formation of the complex MIA ajmaline (Stöckigt and Panjikar, 2007). Ajmaline biosynthesis involves at least eight enzymatic transformations. Five of these enzymes have been purified to homogeneity and putative cDNA sequences were isolated (Ruppert et al., 2005a). Furthermore, three were recombinantly expressed displaying similar properties to the native enzymes (Ruppert et al., 2005a). All the enzymes displayed high substrate specificity delineating the order of the reactions (**Figure 2-3**). The membrane association of the cytochrome P450s has made it difficult to purify these two members (Sarpagan bridge enzyme and vinorine hydroxylase) CYPs involved in ajmaline

biosynthesis and no putative cDNAs have been identified. The last step in the pathway involves the indole nitrogen methylation of norajmaline carried out by an uncharacterized *S*-adenosyl-*L*-methionine (AdoMet)-dependent methyltransferase. Therefore, we decided to focus our efforts to identify the respective cDNAs encoding for remaining steps in ajmaline biosynthesis paying special attention to the AdoMet-dependent methyltransferase catalyzing nitrogen methylation.

### 2.3 *S*-adenosyl-L-methionine (AdoMet)-dependent methyltransferase

AdoMet-dependent methyltransferases (EC 2.1.1.-) are enzymes ubiquitous to all kingdoms of life. This large family of enzymes participates at different levels of primary and secondary metabolism by covalently affecting the hydrophobicity of their substrate, thus regulating its bioactivity and/or bioavailability. AdoMet-dependent methyltransferases (MTs) have diverged from a group of AdoMet-binding proteins that carry out diverse biochemical functions and can be clustered into 15 large superfamilies including MTs (Kozbial and Mushegian, 2005). In this section we will focus on MTs that use small molecules as substrates.

MTs catalyze the transfer of a methyl group from the cofactor AdoMet to an electron rich methyl acceptor atom. This methyl transfer is mostly accomplished through an $S_N2$-like displacement mechanism with nucleophiles attacking the $sp^3$-hybridised methyl carbon of AdoMet (O'Hagan and Schmidberger, 2010). In the active site, the enzyme positions the electron deficient methyl group of AdoMet in close proximity to the chemically reactive nucleophile methyl acceptor atom of the substrate to achieve catalysis, releasing *S*-adenosyl-L-homocysteine (AdoHcy) and the methylated substrate

after completion of the reaction (Liscombe et al., 2012). MTs are commonly classified according to the substrate atom that accepts/abstracts the methyl group, usually good nucleophiles such as oxygen, nitrogen, sulfur and, in some cases, carbon. MTs participating in the biosynthesis of natural products can be grouped into two superfamily fold classes (class I and class III) (Liscombe et al., 2012). To date, all the structurally characterized plant MTs involved in secondary metabolism belong to class I that is characterized by the presence of a "Rossmann-like fold" composed of alternating α-helices and β-strands of the protein, where all the strands are creating a central and relatively planar β-sheet, and the helices are filling the two layers at each side of the plane (Kozbial and Mushegian, 2005; Liscombe et al., 2012).

In all Rossmann-like fold MTs, the AdoMet binding residues occur near the C-terminal end of the alternating β-strands and their adjoining loops (Liscombe et al., 2012). Although residues involved in AdoMet binding are poorly conserved within MTs, three amino acid sequence motifs have been proposed as putative AdoMet binding sites (Joshi and Chiang,1998) based on peptide sequence analyses of 56 functionally characterized plant MTs. Motif A is composed of nine amino acid residues including the highly conserved glycine rich GxGxG signature sequence together with the following consensus amino acids(V/I/L)(V/L)(D/K)(V/I)GGxx(G/A) with close to 100% of them having no more than three mismatches within the consensus sequence (Joshi and Chiang, 1998). This motif spans between the first β-strand and the preceding connecting loop to the α-helix (Kozbial and Mushegian, 2005; Liscombe et al., 2012). The second conserved amino acid motif B (V/I/F)(A/P/E)x(A/P/G)DAxxxK(W/Y/F) in this consensus sequence contains an acidic residue D (Joshi and Chiang, 1998) that usually occurs across the

second β-strand and the adjoining loop (Liscombe et al., 2012). The third conserved amino acid motif C (A/P/G/S)(L/I/V)(A/P/G/S)xx(A/P/G/S)(K/R)(V/I)(E/I)(L/I/V) (Joshi and Chiang, 1998) is part of the third β-strand and is commonly located at the edge of the β-sheet in the overall Rossmann-like fold (Kozbial and Mushegian, 2005). Higher plants contain several MTs which are involved in a wide range of important biological processes such as in decoration of phenolics that are incorporated into lignin (Pinçon et al., 2001), flavonoid/anthocyanin O-methylation related to plant pigmentation (Bouvier et al., 2003), O-methylation of plant hormones that modify their biological activities (Seo et al., 2001), and in the decoration of many other secondary metabolites (De Luca and St. Pierre, 2000).

### 2.3.1 AdoMet-dependent *N*-methyltransferase (NMTs) involve MIA biosynthesis

One of many important biochemical transformations in specialized plant metabolism involves the methylation of nitrogen atoms carried out by NMTs. Such NMTs are involved in the biosynthesis of different classes of alkaloids including purine (Uefuji et al., 2003), benzylisoquinoline (Liscombe and Facchini, 2007), tropane (Hibi et al., 1992) and monoterpenoid indole alkaloids (De Luca et al., 1987). This section will focus on the molecular and biochemical aspects of NMTs involved in MIA biosynthesis.

Pioneering biochemical studies of *N*-methylation in MIA metabolism were carried with protein extracts from *R. serpentina* cell suspension cultures (Stöckigt et al., 1983) and from young *C. roseus* leaves (De Luca et al., 1987). Although these two plant species are specialized in the production of different MIA backbones (**Figure 2-1**), both species contain NMTs that catalyze AdoMet-dependent methylation of the indoline nitrogen of

their respective MIA substrates. The *R. serpentina* NMT involved in ajmaline biosynthesis displayed a strict requirement for the reduced indolenine 1,2- and propenyl 19,20- double bonds from the ajmalan-type MIA substrate (**Figure 2-3**, norajmaline), whereas in the presence of these two double bonds enzyme activity was not detected (Stöckigt et al., 1983). Attempts to purify this norajmaline *N*MT from cell cultures suggested that the enzyme could be associated with membranes (Ruppert et al., 2005a), though no further experiments were performed. The *C. roseus* NMT involved in the 3$^{rd}$ to last step in vindoline biosynthesis also showed strict requirement for the reduced form of the 2,3-alkene located in the aspidosperma terpenoid ring moiety (De Luca et al., 1987). While this NMT activity could not be detected in *C. roseus* cell suspension cultures and in *C. roseus* roots, it was clearly detected in above ground parts of the plant and mostly in young leaves (De Luca et al., 1986, 1987). Subcellular fractionation of organelles by sucrose density gradient centrifugation of isolated broken *C. roseus* leaf protoplasts showed that NMT enzyme activity was located in the chloroplast compartment (De Luca and Cutler, 1987) and associated to the thylakoid membranes (Dethier and De Luca, 1993). Additionally, attempts to purify this membrane-associated *N*MT by using detergents yielded a partially purified soluble enzyme with slight changes in substrate specificity and a putative molecular weight of approximately 60 kDa, further purification steps resulted in significant losses of enzyme activity with no purity (Dethier and De Luca, 1993).

The search for the *C. roseus* NMT gene involved in the vindoline pathway involved a bioinformatics search to mine candidate genes from *C. roseus* expressed sequence tag (ESTs) databases (Liscombe et al, 2010) paying special attention to genes

which contained a chloroplast transit peptides or resembled MTs known to be localized it chloroplasts. The search produced a list of 12 candidate genes containing five ORFs of genes identified as "γ-tocopherol *C*-methyltransferase (γ-TMT)" known to be involved in vitamin E biosynthesis. Candidate genes only expressed in leaves were cloned and expressed in *E. coli* for biochemical characterization and a single recombinant enzyme catalyzed the target reaction that had been identified in earlier studies in C. roseus plants. The recombinant *C. roseus* NMT catalyzed the AdoMet dependent indoline nitrogen methylation of 16-methoxy-2,3-dihydro-3-hydroxy-tabersonine to form 16-methoxy-2,3-dihydro-3-hydroxy-N-methyltabersonine (Liscombe et al., 2010). Phylogenetic analysis of several plant MTs clustered this new *C. roseus* NMT in the same clade as γ-TMTs and separate from other functionally characterized NMTs involved in secondary metabolism (Liscombe et al., 2010). Finally, VIGS of the *NMT* transcripts performed in *C. roseus* seedlings confirmed its biological role in vindoline biosynthesis (Liscombe and O'Connor, 2011) since vindoline levels dropped in silenced lines.

The phytochemical literature has reported the existence of several hundred indole nitrogen methylated MIAs in nature (Buckingham et al., 2010), we hypothesize that members from this newly discovered γ-tocopherol-like *N*-methyltransferase (γ-TLMT) family (Liscombe et al., 2010) are responsible for this functionalization (indole *N*-methylation). Therefore, we decided to search for homologous sequences to the *C. roseus* γ-TLMT in our publicly available large transcriptomic sequencing database http://www.phytometasyn.com/ (Facchini et al., 2012; Xiao et al., 2013) and to select candidate genes for further molecular cloning and biochemical analyses, paying special

attention to transcripts from *R. serpentina*, *Vinca minor* and *C. roseus* for bioinformatics screening.

# Chapter 3
# A picrinine *N*-methyltransferase belongs to a new family of γ-tocopherol-like *N*-methyltransferases (γ-TLMTs) found in medicinal plants that make biologically active monoterpenoid indole alkaloids

**Authors:** Dylan Levac, Paulo Cázares Flores, Fang Yu and Vincenzo De Luca

## 3.1 Introduction

Monoterpenoid indole alkaloids (MIAs) are a large and structurally heterogeneous group of specialized natural products generally derived from the condensation of tryptamine and the monoterpene iridoid, secologanin, to yield the central precursor, strictosidine (**Figure 3-1**) (Salim and De Luca, 2013). The versatility of the strictosidine backbone is revealed by the strictosidine β-glucosidase mediated generation of aglycone intermediates that undergo multiple biochemical rearrangements and substitutions involving cyclases, hydroxylases, oxidoreductases, glycosyltransferases, acyltransferases, methyltransferases, and a few other selected reactions to yield several thousand biologically active MIAs that have been described in the phytochemical literature (Buckingham et al., 2010). MIAs typically occur in the Apocynaceae, Loganiaceae, and Rubiaceae families of plants that have been rich sources of medicines. Drugs from *Rauvolfia serpentina* have been used in treatment of hypertension (ajmalicine) and for diagnostic analyses (ajmaline) while powerful anticancer drugs such as camptothecin from *Camptotheca acuminata* and vinblastine/vincristine from *Catharanthus roseus* continue to be harvested from individual plant species.

In recent years there has been significant progress in identifying new methyltransferase genes and biochemically characterizing their involvement in MIA biosynthesis, with a special emphasis on identifying early and late biochemical steps in

vindoline biosynthesis (Salim and De Luca, 2013; Facchini and De Luca, 2008). Loganic acid *O*-methyltransferase was cloned based on its amino acid sequence similarity to salicylic acid OMT and biochemical characterization of the recombinant enzyme confirmed its role in *O*-methylation of the carboxyl group of loganic acid in the 7$^{th}$ to last step in secologanin biosynthesis (Murata et al., 2008). The 16-hydroxytabersonine *O*-methyltransferase involved in vindoline biosynthesis was biochemically purified, the gene was cloned and its recombinant gene product was functionally characterized to identify the 2$^{nd}$ to last step in vindoline biosynthesis (Levac et al., 2008).

Most recently a new class of methyltransferase, phylogenetically related to γ-tocopherol *C*-methyltransferases was identified and functionally characterized as a 2,3-dihydrotabersonine *N*-methyltransferase [Cr2270 (CrDhtNMT] (Liscombe et al., 2010) (**Table S3-1**; **Figure 3-1**). Prior to the discovery of this novel biochemical function, CrDhtNMT would have been categorized as a γ-tocopherol *C*-methyltransferase and this gave rise to the naming of members of this class of enzyme as γ-tocopherol-like *N*-methyltransferases (γ-TLMT; Liscombe et al., 2010). Additional *in planta* studies using virus induced gene silencing to decrease expression of *CrDhtNMT* suppressed vindoline accumulation in affected leaves in favor of a novel MIA showing the appropriate mass of 16-methoxy-2,3-dihydro-3-hydroxytabersonine (Liscombe et al., 2011) and made a strong case for the involvement of this gene in the 3$^{rd}$ to last step in vindoline and vindorosine biosynthesis. The biochemical localization of CrDhtNMT to chloroplast thylakoids (Dethier and De Luca, 1993) contrasts with the chloroplast envelop localization of tocopherol *C*-methyltransferases involved in vitamin E biosynthesis (Joyard et al., 2009). Also the CrDhtNMT protein may lack a putative chloroplast transit

peptide (Liscombe et al., 2010), raising important questions about the mobilization of the CrDhtNMT gene product to this organelle.

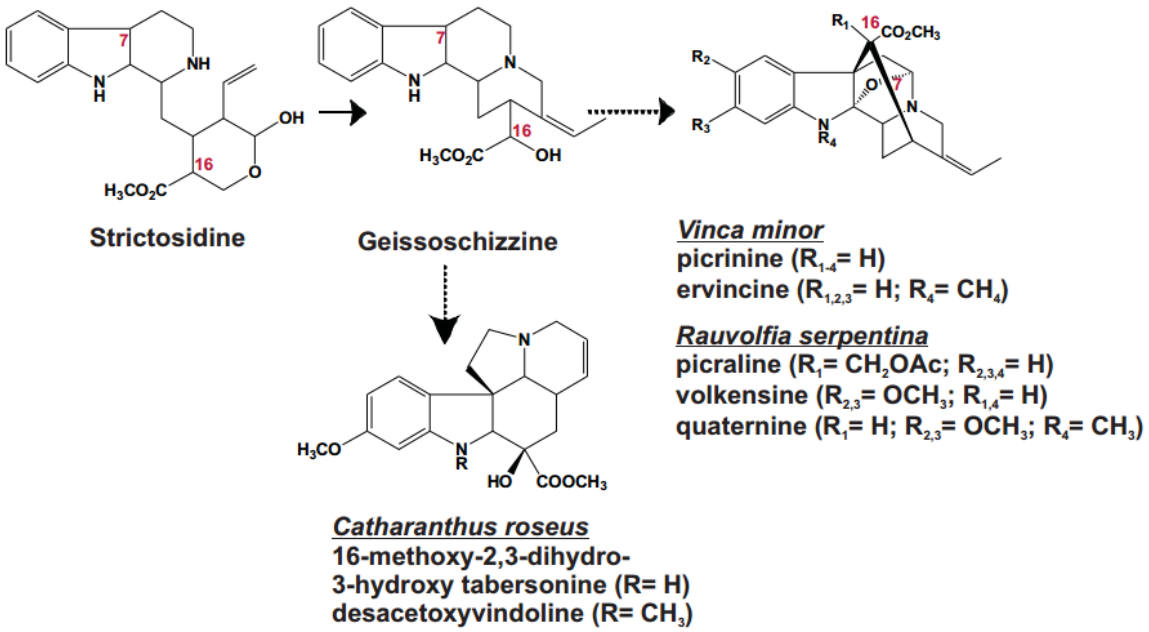



**Figure 3-1: The conversion of strictosidine to various *N*-methylated MIAs in *Catharanthus roseus*, *Rauvolfia serpentina* and *Vinca minor* involves members of a newly discovered family of γ-tocopherol-like *N*-methyltransferase.** The reactions yielding ervincine and desacetoxyvindoline are catalyzed by PiNMT and DhtNMT. The desacetoxyvindoline reaction product of the DhtNMT is then converted to vindoline by desacetoxyvindoline 4-hydroxylase and deacetylvindoline acetyltransferase. Solid arrows represent single enzymatic steps, while dotted arrows represent multiple enzymatic steps. Abbreviations; PiNMT, Picrinine *N*-methyltransferase, DhtNMT, Dihydrotabersonine *N*-methyltransferase.

While several hundred *N*-methylated MIAs that have been described in the phytochemical literature (Buckingham et al., 2010) only CrDhtNMT that catalyses the indoline *N*-methylation of 16-methoxy-2,3-dihydro-3-hydroxytabersonine in the vindoline pathway has been characterized at the biochemical and molecular level. The

present study investigates other γ-TLMTs that may be responsible for similar *N*-methylations of MIAs in other producing plant species. For example picrinine together with a number of picrinine derivatives (picraline, volkensine and quaternine) (**Figure 3-1**) were identified in *Rauvolfia volkensii* (Akinloye et al., 1980a) and in *Rauvolfia oreogiton* (Akinloye et al., 1980b). Picrinine has also been identified in *Vinca minor* (Grossmann et al., 1973) and in *Vinca herbacea* (Boga et al., 2011) while its *N*-methyl derivative, ervincine occurs in *Vinca erecta* (Rakhimov et al., 1967). Candidate γ-TLMT genes were identified by searching annotated EST databases (http://www.phytometasyn.ca/; Facchini et al., 2012; Xiao et al., 2013) of *Catharanthus roseus*, *Vinca minor*, and *Rauvolfia serpentina* for a number of cDNAs that encode putative γ-TLMTs related to CrDhtNMT. We also report the molecular cloning and functional characterization of two novel candidate γ-TLMTs isolated from *Rauvolfia serpentina* (RsPiNMT) and *Vinca minor* (VmPiNMT), respectively, that *N*-methylate picrinine to generate ervincine (**Figure 3-1**).

## 3.2 Results

### 3.2.1 Mining of the *V. minor, R. serpentina* and *C. roseus* databases suggest the presence of a γ-TLMT gene family among members of the Apocynaceae

RNA was extracted from young leaves of *C. roseus* and *V. minor* or from roots of *R. serpentina* and samples were processed for combined 454 sequencing and annotation http://www.phytometasyn.ca (Facchini et al., 2012; Xiao et al., 2013). Mining of each annotated database for γ-TLMTs related to *CrDhtNMT* produced 9 candidate genes that could be investigated for their possible roles as MIA *N*-methyltransferases. A neighbor-joining phylogeny using a class II methyltransferase (*Cr2551*, Liscombe et al., 2010) as an out-group member was constructed together with the 10 γ-TLMTs (Cr91, Cr706,

VmPiNMT, Vm265, Vm2409, Rs820, Rs1755, RsPiNMT, Rs9447, Rs8609; **Table S3-1**), as well as with a previously identified but functionally uncharacterized *Catharanthus* γ-TLMTs (Cr7756; Liscombe et al., 2010) and with γ-known tocopherol MTs (AtGTMT, HaGTMT, BnGTMT, Cr1196) (**Figure S3-1**, **Table S3-1**). The phylogeny was tested by using the UPGMA (Sneath and Sokal, 1973; **Figure S3-1**, A) and Maximum Parsimony (Felsenstein, 1985; **Figure S3-1**, B) methods and bootstrap values greater than 75 are identified. The 2 methods identify a γ-TLMT gene family distinct from that of tocopherol *C*-methyltransferases within members of the Apocynaceae family and the phylogeny based on the UPGMA tree appears to sort MTs according to their biochemical activities clearly separating PiNMTs from other γ-TLMTs and from tocopherol MTs. While the Maximum Parsimony method clusters all γ-TLMTs together, and in a distinct clade from γ-tocopherol C-methyltransferases, it did not resolve PiNMTs from other γ-TLMTs. The added resolution of UPGMA method is supported in the present study and in ongoing studies that remain to be published with other γ-TLMTs where the biochemical function of individual genes is being elucidated.

ClustalW pairwise sequence alignments using the γ-TLMT nucleotide sequences showed that one 1094 bp candidate *Cr91* (365 cluster members; KF896244) shared 99% sequence identity over the open reading frame with the published 16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase [Cr2270, (CrDhtNMT); HM584929] gene isolated from *C. roseus* (cv. Little Bright Eyes) (Liscombe et al., 2010). The predicted open reading frame of *Cr91* consists of 870 bp encoding a protein of Mr = 32 kDa, a putative pI = 6, and a single nucleotide difference predicted to result in a substitution mutation (K145 for E145), when compared to the published *CrDhtNMT*

sequence. We therefore annotated *Cr91* as *CrDhtNMT* and decided to clone this gene in order to conduct comparative substrate specificity studies to those of VmPiNMT and RsPiNMT. Other uncharacterized γ-TLMT-like genes [*Cr2551* (HM584930), *Cr7556*, Liscombe et al., 2010), were not represented in this database. However, *Cr7556* which was identified from a *C. roseus* root EST database (Liscombe et al, 2010) was also present in a previously described *C. roseus* root EST database (KF934425; Murata et al., 2006). The other Catharanthus γ-TLMT from our pyrosequencing database, *Cr706* (KC708453) is not represented in the medicinal plant genomic resource database (medicinalplantgenomics.msu.edu).

ClustalW pairwise amino acid sequence alignment (**Figure S3-1**, **C**) of VmPiNMT (Vm130, 276 cluster members; KC708450) and RsPiNMT (Rs8692, 339 cluster members; KC708448) revealed that they shared 82% amino acid sequence identity (**Figure S3-1**, **D**) and 92% sequence similarity with each other. Remarkably, the uncharacterized *Cr7756* γ-TLMT (Liscombe et al, 2010) appears to be closely related to *R. serpentina* RsPiNMT and *V. minor* VmPiNMT, since it shares 79 % and 88 % amino acid sequence identity to them. In contrast VmPiNMT, RsPiNMT and Cr706 only shared 70%, 72% and 57 % amino acid sequence identity to CrDhNMT (**Figure S3-1**, **D**). This raised the possibility that the transcripts might encode proteins with alternative MIA substrate specificities to those of CrDhNMT.

### 3.2.2 *R. serpentina* and *V. minor* γ-TLMTs are picrinine-*N*-methyltransferases

To identify substrates for the recombinant VmPiNMT and RsPiNMT enzymes MIAs were obtained from *Catharanthus roseus*, *Vinca minor*, *Rauvolfia serpentina*, *Amsonia hubrichtii* and *Tabernamonatana elegans*. Only *Amsonia hubrichtii* extracts appeared to

possess an unknown MIA that could be *N*-methylated in the presence of either recombinant enzyme. Recombinant VmPiNMT [**Figure 3-2**, inset compare lanes 1 (empty vector control) and 2 (recombinant VmPiNMT)] and RsPiNMT [**Figure 3-2**A, inset compare lanes 3 (empty vector control) and 4 (recombinant RsPiNMT)] incubated with *S*-Adenosyl-L-[*methyl*-$^{14}$C]-Methionine ([$^{14}CH_3$]AdoMet) methyl donor and Amsonia extracts converted the unknown MIA into a $^{14}CH_3$-labelled product with an Rf of 5.5. When enzyme assays using non-radioactive AdoMet, MIAs from *A. hubrichtii* and recombinant VmPiNMT were repeated on a larger scale to produce sufficient reaction product for UPLC-MS analysis, a methylated product (possibly ervincine) eluting at 4.25 min. (**Figure 3-2**; VmPINMT + AdoMet) with the expected mass (m/z+ 353) was detected coincident with the loss of an MIA (possibly picrinine) eluting at 3.6 min. (m/z+ 339), while enzyme assays lacking AdoMet showed no methylation of the 3.6 min. peak (**Figure 3-2**; VmPINMT - AdoMet). The mass and absorption spectra of the original MIA (Rt= 3.6 min) and novel (Rt= 4.25 min) methylated MIA peaks suggested that Amsonia surface exudates might contain picrinine that could be converted by these recombinant enzymes to ervincine (**Figure 3-1**). This possible detection of picrinine in *A. hubrichtii* extracts is supported by previous phytochemical analyses that have identified its presence in the related species, *Amsonia sinensis* (Liu et al., 1991).

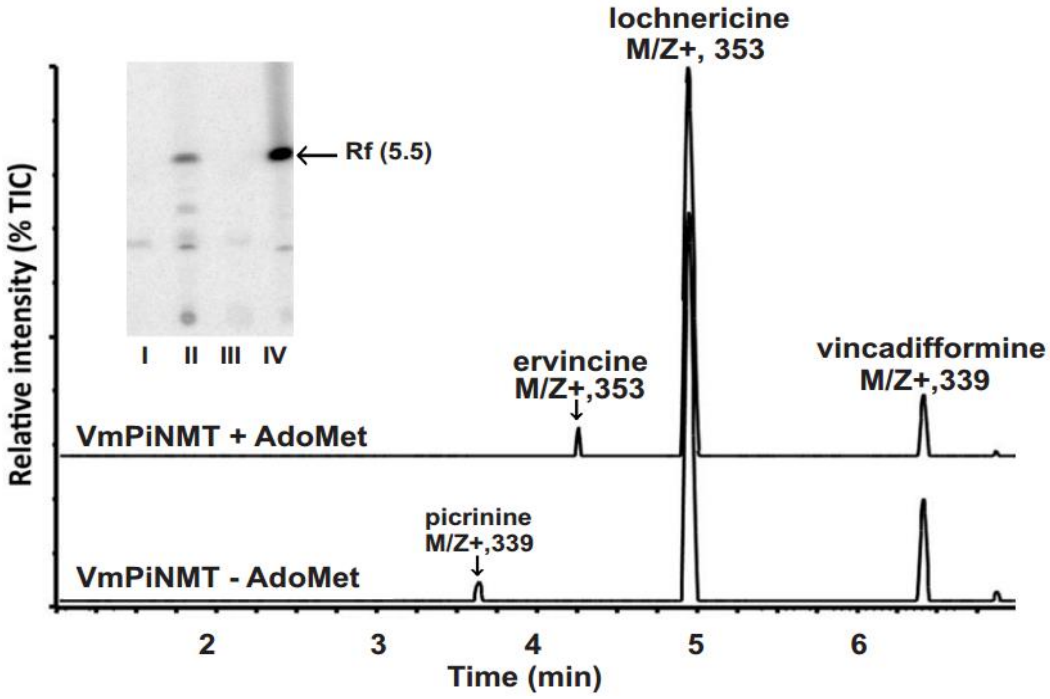



**Figure 3-2: Recombinant VmPiNMT and RsPiNMT enzymes catalyze methyltransferase reactions in the presence of appropriate MIA substrates.** MIAs obtained from leaf surfaces of *Amsonia hubrichtii* were incubated with recombinant VmPiNMT or RsPiNMT enzymes, MIA extracts and [$^{14}CH_3$]-AdoMet as methyl donor. MIA reaction products were submitted to TLC and visualized by autoradiography [**Figure 3-2**, Inset; Lane I (Empty pET 30b vector); Lane II (rVmPiNMT); Lane III (empty pET30b vector); Lane IV (rRsPiNMT)]. Larger scale enzyme assays using rVMPiNMT were then repeated with non-radioactive AdoMet together with larger amounts of *Amsonia hubrichtii* leaf surface MIAs in order to determine the MIA substrate being *N*-methylated as determined by UPLC-MS. Enzyme assays conducted in the presence of VmPiNMT and only Amsonia MIA extract (minus AdoMet) displayed one minor picrinine peak (Rt = 3.62 min, m/z + 339) corresponding to picrinine, together with the major Amsonia MIAs, lochnericine (Rt = 4.9 min, m/z + 353) and vincadifformine (Rt = 6.41 min, m/z + 339). However addition of AdoMet to this assay converted the minor picrinine peak (Rt = 3.62 min, m/z +339) into ervincine (Rt = 4.43 min, m/z + 353).

Larger scale assays with recombinant VmPiNMT (**Figure 3-3A**, upper panel); or RsPiNMT (**Figure 3-3A**, middle panel) enzymes incubated with pure picrinine and with AdoMet clearly converted picrinine (Rt = 3.6 min, m/z+ 339) to ervincine (Rt = 4.25 min, m/z+ 353), compared with enzyme assays without AdoMet that produced no reaction product (**Figure 3-3A**, lower panel). Infrared spectroscopy showed that indole N-H stretching observed between 3250 and 3500 $cm^{-1}$ of picrinine (**Figure 3-3B**, #1) is eliminated in the reaction product (**Figure 3-3B**, #2) confirming that the reaction generates ervincine.

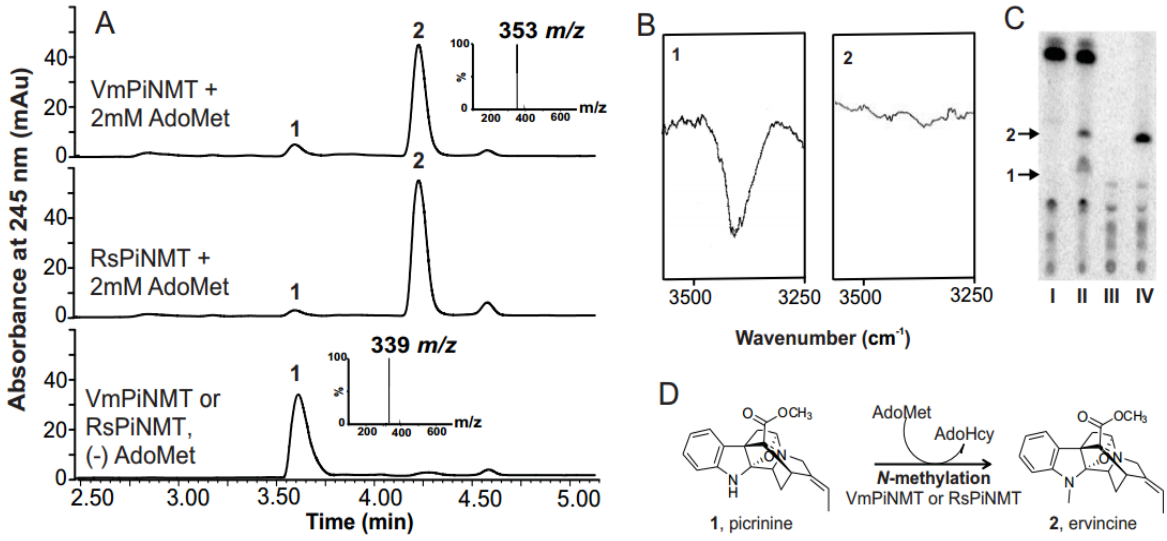



**Figure 3-3: Recombinant VmPiNMT and RsPiNMT catalyze the indole *N*-methylation of picrinine.** UPLC-MS chromatograms show that recombinant VmPiNMT and RsPiNMT will convert picrinine (1: Rt = 3.60, m/z+ 339) into ervincine (2: Rt = 4.25, m/z+ 353) in the presence of 2 mM AdoMet (**A**) compared to enzyme assays without AdoMet or those conducted with the empty pET30b vector. The infrared spectrum of picrinine standard (**B**, #1) shows clear N-H stretching between 3250 and 3500 cm-1 corresponding to the N-H stretch of a secondary amine. The absence of this stretch with the ervincine reaction product (**B**, #2) confirms that the addition of the methyl group takes place on the indole *N*. The presence of PiNMT enzyme activity in

leaves of *V. minor* or *R. serpentina* was shown by performing PiNMT enzyme assays containing crude leaf protein extracts, picrinine substrate and [$^{14}CH_3$]-AdoMet (**Figure 3-3C**). Autoradiograms of reaction products separated by thin layer chromatography show reaction products made by *V. minor* and *R. serpentina* leaf protein extracts assayed in the absence (Lanes 1 and 3) and presence (Lanes 2 and 4) of picrinine. The reaction scheme for the enzymatic conversion of picrinine to ervincine is shown (**Figure 3-3D**).

More detailed substrate specificity studies using a number of MIA substrates (**Tables S2, S3, S4**) showed that while both recombinant VmPiNMT and RsPiNMT preferred picrinine as their best MIA substrate, they also accepted to a minor extent the structurally related 21-hydroxycyclolochnericine and norajmaline (**Table S3-3**) that feature a cyclic ether ring or an ajmalan backbone in their terpene moieties. Among the MIAs that have been described, the akuammilan class of MIAs are produced by cyclization of C-16 to C-7 (such as picrinine), while the akuammidan class of MIAs are produced by cyclization of C-16 to C-5 (such as perivine) (Szabó, 2008; Exkermann and Gaich, 2013). Since neither recombinant enzyme was able to use perivine as a substrate (**Table S3-2** and **Table S3-3**), there may be separate γ-TLMTs that are responsible for producing *N*-methylated derivatives of the akuammidan class of MIAs that have been characterized in the phytochemical literature. The recombinant histidine-tagged RsPiNMT and VmPiNMTs were purified to homogeneity by affinity chromatography (**Figure S3-2**) and were submitted to kinetic analysis. Substrate saturation kinetics for both enzymes produced Kms of 21.0 + 3.0 μM and 20.1 + 4.2 μM for picrinine and 9.3 + 1.1 μM and 15.8 + 2.6 μM for AdoMet, for RsPiNMT and VmPiNMT, respectively (**Figure S3-3**). The Km, Vmax and Kcat/Km values for both recombinant enzymes were

very similar for their respective substrates as those obtained for CrDhNMT (Liscombe et al., 2010; **Figure S3-3**). To complete the basic biochemical characterizations of the two recombinant PiNMTs, their pH optima were determined to be 7.5 and 7.0, for RsPiNMT and VmPiNMT respectively and the temperature optima was 22°C for both enzymes.

### 3.2.3 Picrinine *N*-methyltransferase activity is detected in leaf protein extracts from *Vinca minor* and *Rauvolfia serpentina*

The PiNMT activities of both recombinant enzymes prompted us to test leaf protein extracts from *R. serpentina* and *V. minor* for the presence of PiNMT activities (**Figure 3-3**, **C**). Enzyme assays contained pure picrinine, [$^{14}CH_3$]-AdoMet and crude leaf protein extracts of *V. minor* and *R. serpentina* established the presence of *in planta* PiNMT activity. *V. minor* leaf protein extracts assayed in the absence picrinine produced a major product (Rf = 0.98) and several unknown minor radioactive products (with Rfs < 0.3) (**Figure 3-3**, **C**$_{I}$), while in the presence of picrinine, 2 additional products (Rf, 0.55 and 0.45) (**Figure 3-3, C**$_{II}$) could be seen.

Assays with *R. serpentina* leaf protein extracts displayed similar minor radiolabelled products (with Rfs below 0.3) in the absence of picrinine (**Figure 3-3**, **C**$_{III}$), while a major product (Rf, 0.55) appeared in its presence (**Figure 3-3, C**$_{IV}$). These results indicate that while internal unidentified substrates are being methylated by methyltransferases in crude leaf protein extracts, both plants also contained NMT activities capable of *N*-methylating picrinine.

### 3.2.4 Recombinant *Cr*91 is an authentic CrDhtNMT that does not accept picrinine as a substrate.

Early biochemical characterization of the native enzyme purified from *C. roseus* leaf chloroplasts identified 2, 3-dihydrotabersonine, and its 3-hydroxy derivative to be substrates for *N*-methylation (Dethier and De Luca, 1993). The recombinant enzyme (Liscombe et al, 2010) was shown to catalyze methylation of the indole nitrogen of 2,3 dihydrotabersonine, 2,3-dihytro3-hydroxytabersonine and its 15-methoxyderivatieve by proton and carbon NMR together with high resolution mass spectrometry. In the present study, recombinant Cr91 also catalyzed the turnover of the same MIAs (Liscombe, 2010) as well as 2, 3, 6, 7-tetrahydro-3hydroxytabersonine (**Table S3-4**), confirming it to be the same functionally active CrDhtNMT. However this enzyme did not accept perivine, 21-hydroxycyclolochnericine, norajmaline or numerous other MIAs that were tested (**Table S3-2**).

### 3.2.5 VmPiNMT, and RsPiNMT enzyme activities and transcripts are enriched in organs actively synthesizing MIAs

Crude whole tissue protein extracts were prepared from different organs of *R. serpentina* and *V. minor* as described in methods. Extracts were desalted to remove small molecular weight molecules and were assayed for PiNMT activities using picrinine as substrate. VmPiNMT enzyme activities from *V. minor* 1$^{st}$ leaf pair were 4.5- 22.5- and 9-fold higher compared to those from 2$^{nd}$ leaf pair, 3$^{rd}$ leaf pair and roots respectively (**Figure 3-4, A**). The RsPiNMT activity in *R. serpentina* 1$^{st}$ leaf pair was 1.6-, 2.5-, 12.5- and 2.5-fold higher than those from 2$^{nd}$ leaf pair, 3$^{rd}$ leaf pair, stems and roots, respectively (**Figure 3-4, B**).

The levels of *VmPiNMT* and *RsPiNMT* transcripts were determined by quantitative realtime PCR (RT-qPCR) for the same tissue types analyzed for their respective NMT biochemical activities. The *VmPiNMT* transcript levels were highest in 1$^{st}$ leaf pairs while they were almost not detected in leaf pair 2, leaf pair 3, stems or roots (**Figure 3-4**, **C**). In contrast *RsPiNMT* transcript levels were detected in all organs, but their levels declined continuously with leaf age (**Figure 3-4**, **D**) and were lower in stem or in roots compared to leaf pair 1. These data show that while the enzyme activities found in different tissues (**Figure 3-4, A**-**B**) followed very closely the transcript profiles observed (**Figure 3-4**, **C**-**D**), each gene appeared to be regulated by different developmental and organ specific controls in each plant species. In addition, the biological role(s) of VmPiNMT and RsPINMT remain a mystery since MIA analyses of *V. minor* and *R. serpentina* tissues did not reveal the presence of picrinine or of its *N*-methylated derivative.

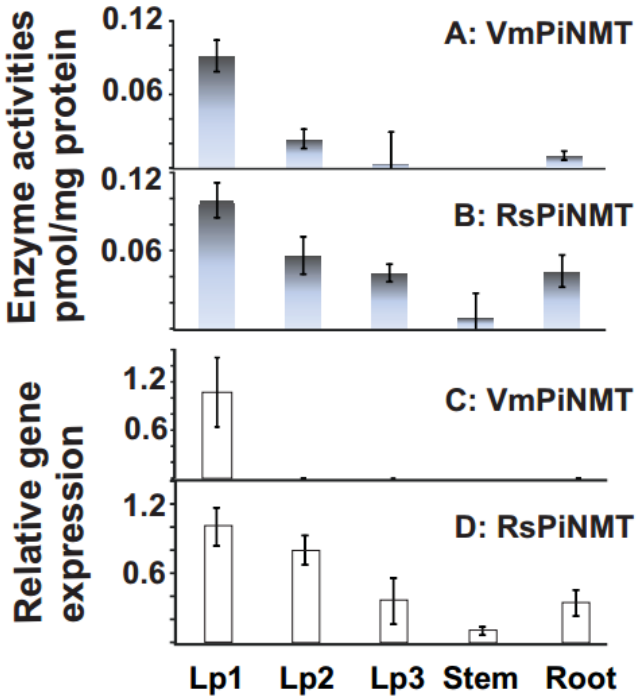



**Figure 3-4: VmPiNMT and RsPiNMT enzyme activities are coordinated with gene expression in different tissues of *V. minor* and *R. serpentina*.** Enzyme activity profiles for VmPiNMT (A) and RSPiNMT (B) in different tissues of each respective plant species. Relative transcript abundance profiles for VmPiNMT (C) and RSPiNMT (D) in different tissues of each respective plant species. Lp1 and Lp2 represents leaf pairs I and 2 next to the leaf apical meristem, respectively.

### 3.2.6 Only the VmTLMT possesses an N-terminal signal peptide

Early investigations into the native CrDhtNMT enzyme employing sucrose gradient fractionation suggested that it is associated with broken and intact chloroplasts, and more specifically, with thylakoid membranes (Dethier and De Luca, 1993). Release of the CrDhtNMT enzyme from thylakoids using detergents showed that it had an apparent Mr of 60,000 as determined by sucrose density gradient centrifugation. Since the cloned CrDhtNMT has a putative subunit Mr of >32,000 (Liscombe et al, 2010), together these data suggest that the functional enzyme may be a dimer (Dethier and De Luca, 1993).

Since tocopherol methyltransferases are known to contain N-terminal plastid transit peptides, it was suggested that the published sequence for *CrDhtNMT* (Liscombe et al, 2010) might not be complete since it lacks this part of the sequence. While the *Cr91-CrDhtNMT* clone has a virtually identical predicted ORF as the *CrDhtNMT* it has an additional 224 nucleotides upstream of the Kozak consensus sequence. These results suggest that *CrDhtNMT* may not possess an amino-terminal chloroplast transit peptide, although this needs to be further investigated. Remarkably, ClustalW multiple amino acid sequence alignment of the CrDhtNMT, VmPiNMT, and RsPiNMT revealed highly similar methyltransferase domains absent of gaps (**Figure S3-1**, **C**), while WolfP and TargetP 1.1 analysis of each protein predicted that only VmPiNMT contained a clearly identified 25 amino acid residue for a putative secretory transit peptide that may target this enzyme to extracellular spaces (Horton et al., 2007). The predicted open reading frame of *VmPiNMT* (969 bp) encodes a 35 kDa protein (pI = 5.84) that when processed changes its Mr to 32 kDa and reduces its apparent pI to 5.39. The *RsPiNMT* (873 bp) also encodes a 32 kDa protein (apparent pI of 6.9).

### 3.3 Discussion

A novel class of tocopherol like *N*-methyltransferase involved in the 3rd to last step in vindoline biosynthesis was recently discovered (Liscombe et al., 2010; Liscombe and O'Connor, 2011) and its sequence was used in the present study for large scale transcriptome analysis of several MIA producing plant species (www.phytometasyn.ca, Facchini et al., 2012; Xiao et al, 2013). This study identifies 9 new γ-TLMTs (**Figure S3-1**) and reports the biochemical function of a second member of this emerging family that is involved in the *N*-methylation of picrinine in *V. minor* (VmPiNMT) and *R. serpentina*

(RsPiNMT). In addition the evolution of different γ-TLMTs with different *N*-methylation substrate specificities may have led to the biosynthesis of at least 150 known MIAs (Buckingham et al, 2010) that have been discovered by phytochemists and that could be responsible for taxonomical clustering of indole *N*-methylated MIAs within the Apocynaceae.

The discovery of VmPiNMT and RsPiNMT was in part based on an interesting strategy for identifying the biochemical functions of different members of the γ-TLMT class of enzymes by harvesting potential MIA substrates from the leaf surfaces of plants that accumulate them and using them for biochemical activity testing (**Figure 3-2**). The method depends on sufficient amounts of substrate being available in the extract for conversion into product that can be detected in radioactive methyltransferase assays, followed by larger scale detection of products by UPLC MS (**Figure 3-2**). The use of radioactive assays (**Figure 3-2**) was essential for preliminary screening of MIA metabolites from leaf surface extracts and this led to the identification of the minor MIA, picrinine that acts as *N*-methyl acceptor (**Figure 3-2**). While it was fortuitous that picrinine was commercially available for biochemical verification of VmPiNMT and RsPiNMT function, this methyl acceptor could have been purified from the source tissue for subsequent use in more detailed enzyme characterization of the recombinant proteins. This strategy to identify gene function is particularly useful in non-model plants systems, like *R. serpentina* and *V. minor*, where transient or stable gene knock-down methodologies have not been developed, or where these gene knock-down strategies do not yield clear target molecules because they may be unstable or they may be redirected to alternative pathways.

Substrate specificity studies showed that VmPiNMT and RsPiNMT *N*-methylated picrinine and to a much lower extent of 21-cyclolochnericine and norajmaline (**Table S3-3**); they did not *N*-methylate tabersonine-related substrates that were accepted by recombinant CrDhtNMT (**Table S3-2** and **Table S3-4**) and they were unable to use numerous other MIAs that were tested (**Table S3-2**). Similarly CrDhtNMT did not *N*-methylate picrinine, 21-cyclolochnericine or norajmaline (**Table S3-4**).

Previous studies with *C. roseus* CrDhtNMT have suggested that this enzyme is associated with chloroplast membranes and experiments with the detergent solubilized enzyme established that it behaves as a 60 kDa dimer (Dethier and De Luca, 1993) and as an undefined larger molecular weight aggregate in the absence of detergent. The lack of a transit peptide in CrDhtNMT raises questions about the mechanism involved to mobilize this enzyme to the chloroplast compartment. In contrast VmPiNMT does have a putative secretory transit peptide that raises questions about its secretory role that can only be defined by detailed biochemical and cell biology localization studies.

The recombinant RsPiNMT, VmPiNMT enzymes (**Figure S3**) have very similar properties to those of CrDhtNMT (Liscombe et al., 2010) in terms of their similar Km, Vmax and Kcat/Km values for their respective substrates. The inability of either RsPINMT or VmPiNMT to *N*-methylate perivine or other MIAs provides support that other substrate specific γ-TLMT family members may be involved in the biosynthesis of some of the 150 known *N*-methylated MIAs (Buckingham et al., 2010). While RsPINMT did convert norajmaline to ajmaline we do suspect that a separate γ-TLMT is likely to be responsible for the biosynthesis of ajmaline. While picrinine or its *N*-methylated derivative were not detected in our test *V. minor* or *R. serpentina* plants, it was previously

detected in *V. minor* (Grossmann et al., 1973; Buckingham et al., 2010) and in *Vinca herbacea* (Boga et al., 2011), whereas ervincine was also detected in *Vinca erecta* (Rakhimov et al., 1967; Liu et al, 1991) (**Figure 3-1**). In addition picrinine, picraline, volkensine and quaternine (**Figure 3-1**) accumulate in *Rauvolfia volkensii* (Akinloye et al., 1980a) and *Rauvolfia oreogiton* (Akinloye et al., 1980b). Members of the genus *Alstonia* produce a range of picrinine (Hai et al, 2008) and picraline (Arai et al., 2010) derivatives that include powerful *N*-methylated drug candidates that inhibit the Na+-glucose cotransporter involved in the re-absorption of glucose in human kidneys. In addition, picrinine from *Alstonia scholaris* (Chatterjee et al., 1965) has been documented to have powerful antitussive, anti-asthmatic and expectorant properties (Shang et al., 2010). We conclude that the many *N*-methylated derivatives of picrinine found in Nature are likely to involve the new PINMT genes described in the present study.

## 3.4 Materials and methods

### 3.4.1 Plant material

*Catharanthus roseus* (*cv* Little Delicata), *Rauvolfia serpentina*, and *Vinca minor* plants were grown in a greenhouse under a long-day photoperiod (16:8 hrs) at 28$^{o}$C. Young leaves (< 1.5 cm in length), whole root, and first leaf pair respectively were harvested for isolation of total RNA and for use in next generation 454 Roche pyrosequencing. Other tissues to be used for other experiments were also harvested and extracted in the same manner.

### 3.4.2 MIA isolation from leaf surfaces

Fresh leaves (approximately 10 g) from *Amsonia hubrictii*, *Catharanthus roseus*, *Vinca minor* and *Rauvolfia serpentina*, were submerged in chloroform in Erlenmeyer flasks and shaken at 100 RPM for 2 h in order to extract surface MIAs from each plant species. The chloroform was decanted into a round bottom flask and extracts were taken to dryness by flash evaporation (Buchi Rotavapor R-205, Heating Bath B-490, Brinkmann WLK 230 LAUDA, Thermosavant GP 110 GelPump). The residue was dissolved in methanol and water (85:15 water:methanol) acidified to pH 3 with 10% $H_2SO_4$. Contaminating small molecules were extracted into ethyl acetate that was discarded and 10 N NaOH was added to the aqueous phase (pH 12). MIAs were extracted into ethyl acetate and this crude MIA fraction was taken to dryness by flash evaporation, dissolved in methanol and stored at -20$^{o}$C until use for enzyme assays.

### 3.4.3 Total RNA isolation

Plant tissues were harvested from the greenhouse and homogenized in a mortar and pestle with liquid nitrogen to produce a fine powder. Approximately 0.2 g of powdered homogenate was transferred to 1.5 mL microcentrifuge tubes containing 0.225 mL extraction buffer (0.1 M NaCl, 0.01 M Tris-Cl pH 7.5, 1 mM EDTA, 1% SDS) and 0.15 mL phenol:chloroform:isoamyl alcohol. Samples were then shaken vigorously for 15 minutes at room temperature, and centrifuged (9500 g, 10 minutes). The aqueous phase was transferred to fresh microcentrifuge tubes and 0.1 volume 3 mM NaOAc pH 5.1, and 2 volumes cold 95% ethanol was added to precipitate nucleic acids over 20 minutes at -80 $^{o}$C. Precipitated nucleic acids were pelleted by centrifugation (9500 g, 15 minutes, 4 $^{o}$C). The supernatant was removed, the nucleic acid pellets were washed with 70% ethanol

and after centrifugation (10,000 xg, 10 minutes, 4 $^{o}$C) the ethanol was removed and pellets were dried using a SAVANT DNA120 (Thermo Fisher). Dry pellets were dissolved in 0.5 mL sterile DEPC treated water, and 1 volume 4 M cold LiCl was added with gentle mixing and incubated on ice for 3 hours. Nucleic acids were pelleted by centrifugation (10,000 xg, 10 minutes, 4 $^{o}$C), the supernatant was removed and pellets were dissolved in sterile water. Samples were treated with DNase for 30 minutes at 37 $^{o}$C to remove genomic DNA and purified total RNA was mixed with 0.1 volume 3 mM NaOAc pH 5.1, and 4 volumes 100% ethanol and centrifuged (10,000 xg, 10 minutes, 4 $^{o}$C). The pellet was washed three times with 75% ethanol to remove salts, and stored in 100% ethanol at 80$^{o}$ C until needed.

### 3.4.4 Sequencing

Databases enriched in sequences encoding genes involved in MIA biosynthesis were generated by extracting RNA isolated from highly active MIA producing tissues (*C. roseus* young leaves, *R. serpentina* whole root, and *V. minor* 1$^{st}$ leaf pair). The RNA was provided to the McGill University Genome Quebec Innovation Centre (Montreal, QC, Canada) for processing and submission to pyrosequencing on a Genome Sequencer FLX (454 Life Sciences Corp., Bradford, CT). Sequence fidelity was verified by traditional Sanger sequencing at Genome Quebec Innovation Center for all PCR products used to establish gene expression profiles or for cloning open reading frames (ORFs) to be used for biochemical characterization of recombinant enzymes.

### 3.4.5 Database assembly and transcript annotation

The data obtained from pyrosequencing of *C. roseus*, *R. serpentina* and *V. minor* materials was assembled using our bioinformatics pipeline as described previously (Facchini et al., 2012; Xiao et al., 2013). The databases are publicly available and can be found at www.phytometasyn.ca

### 3.4.6 Database mining and identification of γ-TLMTs

The nucleotide sequence of CrdhTNMT (Liscombe et al., 2010) was used in a blast search to identify other γ-TLMTs in the *C. roseus*, *R. serpentina* and *V. minor* and phylogenetic analysis was used to resolve γ-TLMTs from authentic γ-tocopherol MTs.

### 3.4.7 Phylogenetic analysis

#### 3.4.7.1 Amino-Acid alignment of tocopherol and tocopherol-like methyltransferases

Tocopherol and Tocopherol-like peptide amino acid sequences (**Table S3-1**) were copied into the Molecular evolutionary Genetics Analysis version 5 application (MEGA 6) integrated tool for phylogenetic analysis (Tamura et al., 2013), and peptide sequences were aligned using Multiple Sequence Comparison by Log-Expectation (MuSCLE) (Edgar, 2004) with the following constraints; Gap Penalties: Open = -2.9, Extend = 0, Hydrophobicity Multiplier = 1.2. Memory/Iterations: Max Memory in Mb = 4095, Max iterations = 50. Clustering Method UPGMB.

#### 3.4.7.2 Phylogeny construction

The evolutionary history was inferred with two different methods. Using the UPGMA method (Sneath and Sokal, 1973) the bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein

1985). Branches corresponding to partitions reproduced in less than 75% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (Felsenstein 1985). The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved 13 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 275 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al 2013). Using the Maximum Parsimony method the bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein 1985). Branches corresponding to partitions reproduced in less than 75% bootstrap replicates are collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches (Felsenstein 1985). The MP tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (Niel and Kumar, 2000) with search level 1 in which the initial trees were obtained by the random addition of sequences (10 replicates). The analysis involved 13 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 275 positions in the final dataset.

Evolutionary analyses were conducted in MEGA6 (Tamura et al 2013).

### 3.4.8 Molecular cloning of *N*MTs from *C. roseus*, *V. minor* and *R. serpentina*

Full-length cDNA sequences for *CrDhtNMT* (HM584929), *Vm130* (*VmPiNMT; A*ccession number KC708450), and *Rs8692* (*RsPiNMT*; Accession number KC708448)

were identified in our *C. roseus* cv. little delicata, *V. minor*, and *R. serpentina* databases respectively by performing a local BLASTp using the published amino acid sequence of CrDhtNMT as query. PCR primers were designed to amplify the known *CrDhtNMT* open reading frame (ORF) (DhtNMT_FW 5'-TTCATATGGAAGAGAAGCAGGAG-3', DhtNMT_Rv 5'-TTGCGGCCGCATATTGATTTTCGTCCGTAAC-3'), the putative *VmPiNMT* ORF lacking signal peptide, (VmPiNMT_Fw 5'-TTCATATGGCGGAAAAGCAAG-3', VmPiNMT_Rv 5'-TTGCGGCCGCATTTAGATTTGCGGCATGTAAC-3'), the putative *VmPiNMT* ORF with signal peptide (S_VmPiNMT_Fw 5'-TTCATATGTACACTTGTTCAATTATAATATATAT-3', S_VmPiNMT_Rv 5'-TTGCGGCCGCATTTAGATTTGCGGCATGTAAC-3') and the putative *RsPiNMT* ORF (RsPiNMT_Fw 5'-TTCATATGGCAGAGAAGCAGCAGGC-3', RsPiNMT 5'-TTGCGGCCGCATTTTGATTTCCTGCATGTAATTGCAAC-3'). Note the presence of 5' NdeI and 3' NotI restriction sites. These restriction sites were incorporated to facilitate directional cloning into the pET30b (Invitrogen) *E.coli* expression vector. Total RNA from each species was obtained as described in earlier sections and cDNA synthesis was performed using AMVreverse Transcriptase (Promega). Each cDNA pool was used to clone each desired ORF by PCR with Phusion High-Fidelity DNA polymerase (New England Biolabs) using the following conditions; 2 minute initial denaturation, 95 $^{o}$C; 35 cycles of (20 second, 95 $^{o}$C; 20 seconds, 57 °C; 60 seconds 72 °C) followed by a 5 minute final extension at 72 °C. PCR products were then precipitated and washed with ethanol, dissolved in MilliQ water, and used directly for A-tailing reactions (The reaction was conducted with ExTaq DNA polymerase (TaKaRa) under standard conditions in the

presence of dATPs for 1 hour at 72$^{o}$C). A-tailed PCR products were used directly for ligation into pGEM-T Easy T-A cloning vector (Promega). The ligation reaction was then transformed into XL-10 Gold ultracompetent *E.coli* cells, and plated on LB ampicillin (100 μg/mL), agar plates with XGAL and IPTG to facilitate blue-white colony selection after being grown overnight at 37$^{o}$C. White colonies from each plate were screened by colony PCR to verify presence of inserts and verified by sequencing (McGill University Genome Quebec DNA sequencing services, Montreal, Quebec). One colony was then selected for large scale plasmid purification, NdeI-NotI double digestion, and directional cloning into pET30b *E.coli* expression vector. Presence of desired ORF in pET30b, and proper direction was verified by PCR using a combination of gene specific and vector specific PCR primers.

### 3.4.9 Production and purification of the recombinant *VmPiNMT* and *RsPiNMT*

*Escherichia coli* BL21-codonPlus (DE3)-RIL (Stratagene) harboring the pET30-b/VmPiNMT or pET30-b/RsPiNMT were cultivated in 1 L of Luria-Bertani medium supplemented with 50 mg $L^{1}$ kanamycin at 37 $^{0}$C with shaking (250 rpm) to an $OD_{600}$ of 0.6, and then induced with IPTG (1 mM final concentration) at 20 $^{0}$C with shaking (125 rpm). After 9 h of growth at 20 $^{0}$C, cells were collected by centrifugation (3000 *g*) for 15 min at 4 $^{0}$C. Pelleted cells were resuspended in buffer A [100 mM Tris-HCl pH (7.7), 300 mM NaCl, 10% glycerol, 14 mM β-mercaptoethanol, 10 mM imidazole] and lysed by sonication. Cell debris was removed by centrifugation (3,000 x *g*) for 25 min at 4 $^{0}$C. The supernatant was submitted to a Ni-NTA Agarose (Qiagen) column chromatography (previously equilibrated in buffer A), which was then washed in a stepwise manner with buffer A containing increasing concentration (20, 50, 100, and 250 mM) of imidazole.

Recombinant RsPiNMT was purified by elution with 20 mM imidazole while recombinant VmPiNMT was purified by elution with 100 mM imidazole elution (**Figure S3-2**). To remove the imidazole, PiNMTs were desalted using a PD Spin Trap G-25 column (GE Healthcare) equilibrated with buffer B [100 mM Tris-HCl pH (7.7), 14 mM β-mercaptoethanol].

### 3.4.10 Extraction of enzyme activities from *C. roseus, V. minor* and *R. serpentina* leaves

First leaf pairs (5 g) were harvested from *C. roseus, V. minor,* and *R. serpentina* to be harvested and homogenized in a cooled mortar and pestle with 15 mL ice cold 100 mM Tris-Cl pH 7.7, 14 mM mercaptoethanol. Leaf homogenates were then filtered through cheese cloth into a 15 mL conical tube and filtrates were centrifuged at 1000 x g to remove large debris. The resulting supernatant was desalted by Sephadex G25 chromatography, and used directly for enzyme assays.

### 3.4.11 NMT Enzyme assays

NMT activities were determined using a radioactive enzyme assay (150 µL) containing protein (recombinant or from plant extracts), 2.5 nCi *S*- ([$^{14}CH_3$])-AdoMet (specific activity 58 mCi/mmol; GE Healthcare Canada), and 5 µg substrate (2,3-dihydro-3hydroxytabersonine for DhtNMT, or picrinine for the 2 PiNMTs), 100 mM Tris-HCl, pH 7.7, 14 mM mercaptoethanol). Enzyme assays were incubated at 37$^{o}$C for 1 hour. Assays were stopped by adding 10% v/v 10N NaOH, and reaction products were extracted to 500 µl ethyl acetate. 50 µL of the reaction products was used for quantification by scintillation (Beckman LS 6500 Scintillation Counter). The standard nonradioactive enzyme assay (200 µL) contained crude desalted recombinant enzyme, 2

mM AdoMet, 5 µg substrate, and enzyme assay buffer (100 mM Tris-HCl, pH 7.7, 14 mM mercaptoethanol). Enzyme assays were incubated at 37$^{o}$C for 3 hours.

Assays were stopped by adding 10% v/v 10N NaOH and reaction products were extracted to 500 µl ethyl acetate. Organic phases of 3 replicate assays were combined into a single 2 mL tube, and taken to dryness using a SAVANT Speedvac (Thermo Scientific). Dried alkaloid products were re-suspended in 200 µL methanol, and filtered through a 0.22 µm PALL filter (VWR Canada).

Reaction products were analyzed by ultra-high performance liquid chromatography with diode array detections and mass spectrometry UPLC-MS (Waters) where MIAs were separated using an Aquity UPLC BEH C18 column with a particle size of 1.7 µm and column dimensions of 1.0 × 50 mm. Samples were maintained at 4 °C and 5-µL injections were made into the column. The solvent systems for MIA analyses were as follows: solvent A, methanol: acetonitrile:5-mM ammonium acetate and 6:14:80; solvent B, methanol: acetonitrile:5-mM ammonium acetate at 25:65:10. The following linear elution gradient was used: 0–0.5 min 99% A, 1% B at 0.3 mL/min; 0.5–0.6 min 99% A, 1% B at 0.4 mL/min; 0.6–7.0 min 1% A, 99% B at 0.4 mL/min; 7.0–8.0 min 1% A, 99% B at 0.4 mL/min; 8.0–8.3 min 99% A, 1% B at 0.4 mL/min; 8.3–8.5 min 99% A, 1% B at 0.3 mL/min; and 8.5–10.0 min 99% A, 1% B at 0.3 mL/min. The mass spectrometer was operated with a capillary voltage of 2.5 kV, cone voltage of 34 V, cone gas flow of 2 L/h, desolvation gas flow of 460 L/h, desolvation temperature of 400 °C, and a source temperature of 150 °C.

### 3.4.12 Substrate specificity assays

Desalted crude recombinant protein extracts (r*PiNMT*, r*DhtNMT*) were used to assay pure alkaloid substrates from our collection to identify small molecules accepted by each NMT. Initial screening used the standard radioactive enzyme assay with an incubation time of 3 hours. Any substrates that yielded greater than 3x background radioactive counts were repeated using nonradioactive assay conditions in batches of 3 replicates. Replicates of nonradioactive enzyme assay products were extracted to organic phase, pooled, dried and prepared for UPLC-MS analysis as described previously.

### 3.4.13 Picrinine *N*MT Kinetic analysis – Saturation curves

Picrinine *N*MT assays contained 700 ng of affinity purified protein (**Figure S3-2**), 0.9-80 µM Picrinine, 1.35-50 µM AdoMet [mixed with 0.006 µCi ([$^{14}CH_3$])-AdoMet 56.3 mCi mmol$^{-1}$ specific activity (Perkin-Elmer)], in a total volume of 50 µL. Kinetic assays were performed under optimal conditions (30 $^{0}$C, buffer B), incubated for 15 min and terminated with the addition of 15 µL of 10 M NaOH. The reaction product was extracted with 400 µL ethyl acetate and centrifuged at 10,000 x *g* for 1 min. Finally, 200 µL of the organic phase was analyzed using a Beckman LS 6500 Scintillation counter. For picrinine kinetics, AdoMet was fixed at 100 µM [mixed with 0.014 µCi ([$^{14}CH_3$])-AdoMet 56.3 mCi mmol$^{-1}$ specific activity (Perkin-Elmer)], and for AdoMet kinetics picrinine was fixed at 200 µM. Kinetic data were fitted by non-linear regression and analyzed using GraphPad Prism version 5 for Windows (GraphPad Software).

### 3.4.14 Reverse-Transcription quantitative PCR (RT-qPCR)

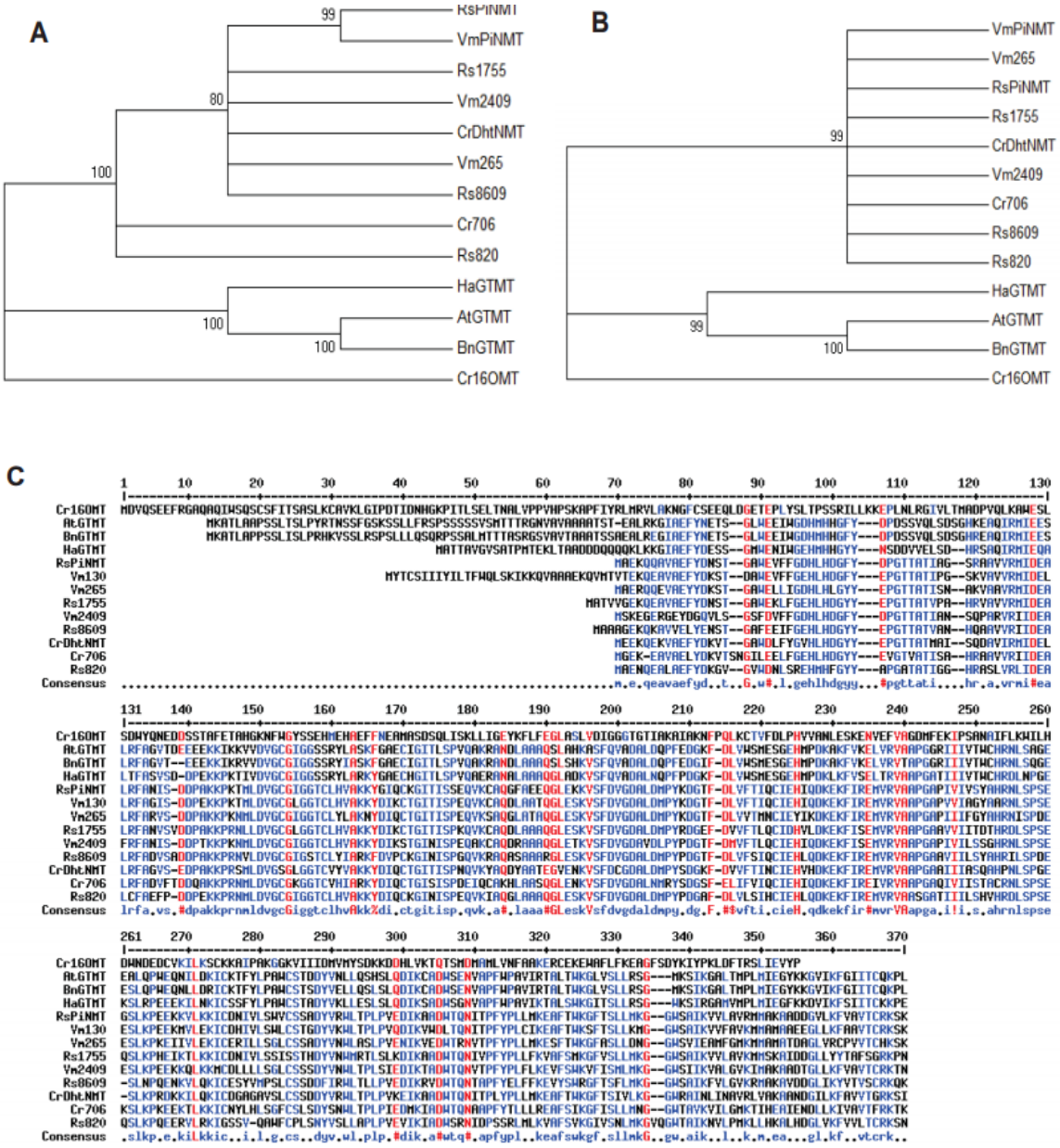
Gene expression was monitored by **RT-qPCR** using PCR primers flanking the stop codon such that the 3' UTR of the transcript assists in increasing the specificity of the PCR reaction. The primer pairs Actin-F (5'-GGAGCTGAGAGATTCCGTTG-3') and Actin-R (5'-GAATTCCTGCAGCTTCCATC-3'), VmPiNMT-F (5'-TTCTCTCATGGCTTTGTTCT-3') and ViPiNMT-R (5'-TTAACGAAATTTTTGGCATT-3'), RsPiNMT-F (5'-ACATGCAGGAAATCCAAATA -3') and RsPiNMT-R (5'-ATTCATCAAGAGCTCCACAC -3') were used to generate 73 bp, 274 bp and 302 bp length PCR products respectively. The RT-qPCR reaction was carried out in a final 25 µl containing 200 nM of each primer, 12 µl iQ$^{TM}$ SYBR Green PCR Master Mix (Bio-Rad), and 1 µl of cDNA (corresponding to approximately 286 ng cDNA) using the following conditions: 95$^{o}$C for 15 min, then 40 cycles of 95$^{o}$C for 10 sec, 55$^{o}$C for 15 sec and 72$^{o}$C for 30 sec using the CFX96 real-time SYBR system (Bio-Rad). Three biological replicates were collected for each *R. serpentina*, and *V. minor* tissue and single qPCR reactions were conducted for each biological replicate.

## 3.5 Acknowledgements

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (V.D.L.), Canada Research Chairs (V.D.L.), Genome Canada, Genome Alberta, Genome Prairie, Genome British Columbia, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada and other government and private sector partners. We recognize the skilled technical work of next-generation sequencing personnel at the

McGill University-Genome Québec Innovation Centre. We are grateful to Christoph Sensen, Mei Xiao and Ye Zhang for their dedicated bioinformatics support and large scale gene annotation efforts that that helped in the identification of γ-TLMTs from the Phytometasyn web site. Paulo Cázares was supported by a postgraduate PhD scholarship from Consejo Nacional de Ciencia y Tecnología, México (CONACyT). We would like to thank the Mr. Lorne Fast, Curator of Collections, Niagara Parks Botanical Gardens and School of Horticulture, for provision of *Amsonia hubrichtii* plants used in part of this study.

## 3.6 Supporting information



D

| Amino Acid sequence identities between member TLMTs | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | VmPiNMT | Vm265 | Vm2409 | RsPiNMT | Rs820 | Rs1775 | Rs8609 | CrDhtNMT | Cr706 |
| VmPiNMT | 100 | | | | | | | | |
| Vm265 | 72 | 100 | | | | | | | |
| Vm2409 | 73 | 64 | 100 | | | | | | |
| RsPiNMT | 82 | 72 | 77 | 100 | | | | | |
| Rs820 | 61 | 59 | 60 | 65 | 100 | | | | |
| Rs1755 | 73 | 66 | 66 | 74 | 61 | 100 | | | |
| Rs8609 | 65 | 63 | 67 | 70 | 61 | 67 | 100 | | |
| CrDhtNMT | 70 | 63 | 63 | 72 | 57 | 66 | 63 | 100 | |
| Cr706 | 63 | 59 | 63 | 65 | 59 | 64 | 61 | 57 | 100 |

**Figure S3-1: Comparative phylogenetic and amino acid analyses of various tocopherol MTs and tocopherol-like NMTs.** The phylogeny was tested by the **A**) UPGMA (Sneath and Sokal, 1973) and by the **B**) Maximum Parsimony (Felsenstein, 1985) methods using tocopherol, and tocopherol-like methyltransferases constructed from peptide sequences listed in **Table S3-1** and aligned using MuSCLE. *Cr2551*, a Class II methyltransferase (Liscombe et al. 2010) was used as outgroup. **C**) ClustalW multiple amino acid sequence alignment of *Cr16OMT*, *AtGTMT*, *BnGTMT*, *HaGTMT*, *Rs8692* (*RsPiNMT*), *Vm130* (*VmPiNMT*), *Vm265*, *Rs1755*, *Vm2409*, *CrDhtNMT* and *Cr706* showing high conservation of the γ-TLMT methyltransferase domain, and a unique N-terminal predicted secretory peptide for only *VmPiNMT*. Neither *RsPiNMT* nor *CrDhtNMT* contain N-terminal targeting peptides. **D**) Amino acid identities (%) between tocopherol-like *N*MTs.

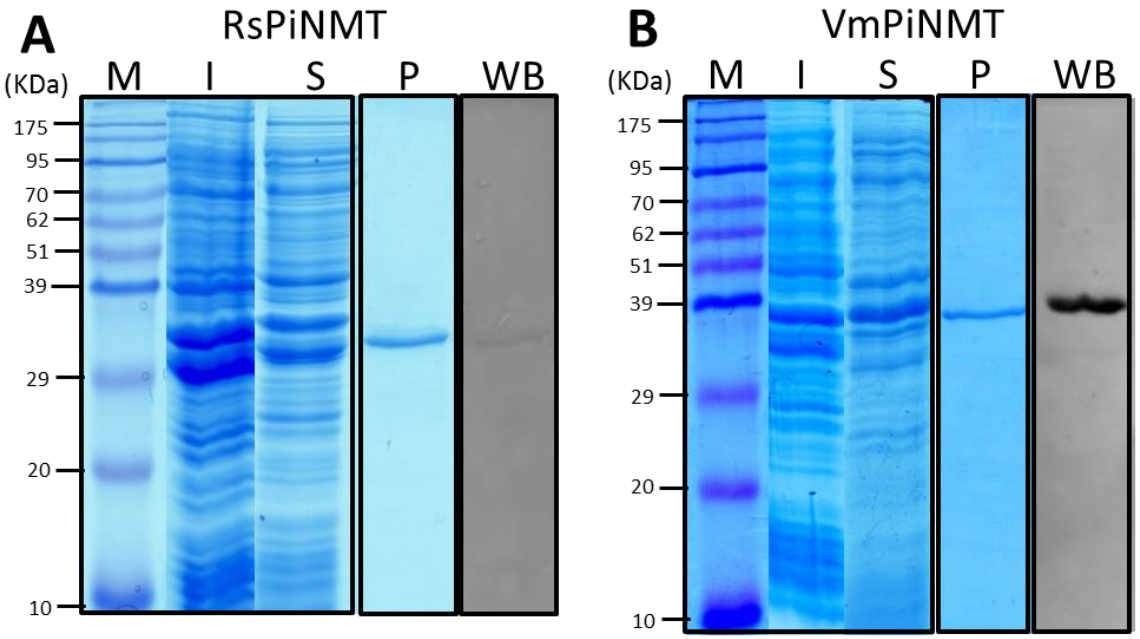
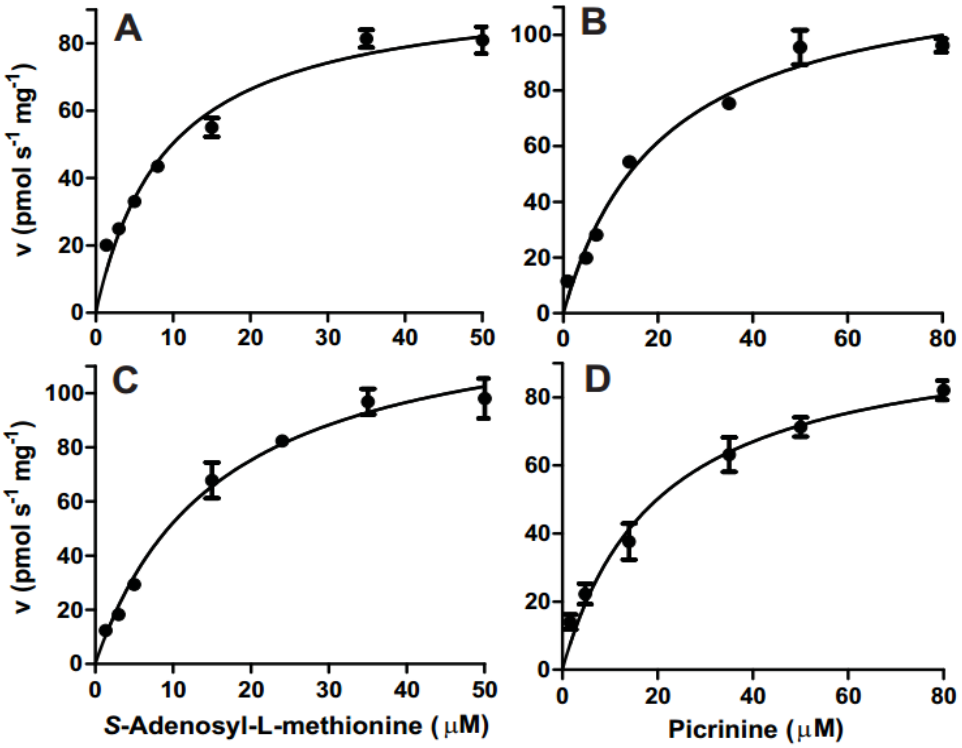



**Figure S3-2: SDS-PAGE analysis of purified recombinant proteins.** (**A**) RsPiNMT and (**B**) VmPiNMT. Lanes are labeled as follows: M, marker; I, crude insoluble protein; S, crude soluble protein; P, His-tag purified protein; and WB, western blot.




| Enzyme | Substrates | $K_m$ (μM) | $V_{max}$ (pmol $s^{-1}$ $mg^{-1}$ protein) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($s^{-1}$ $M^{-1}$) |
|---|---|---|---|---|---|
| RsPiNMT | AdoMet | 9.3 ± 1.1 | 97.2 ± 4.1 | 4.5 ± 0.1 | 483,871 |
| | Picrinine | 21.0 ± 3.0 | 126.1 ± 6.6 | 5.8 ± 0.3 | 276,190 |
| VmPiNMT | AdoMet | 15.8 ± 2.6 | 135.0 ± 8.6 | 6.9 ± 0.4 | 436,709 |
| | Picrinine | 20.1 ± 4.2 | 100.8 ± 7.2 | 5.1 ± 0.3 | 252,725 |

**Figure S3-3: Michaelis-Menten plots of kinetic data for purified, recombinant picrinine *N*-methyltransferases.** Non-linear regression analysis of kinetic data were used to estimate steady-state kinetic parameters for *RsPiNMT* with substrates (**A**) S-adenosyl-L-methionine (AdoMet), (**B**) picrinine; and for *VmPiNMT* with substrates (**C**) AdoMet and (**D**) picrinine. Table: Steady-state kinetic parameters for r*RsPiNMT* and r*VmPiNMT* recombinant enzymes for Picrinine and AdoMet at saturating alternative substrate concentrations.

**Table S3-1: Abbreviations, Annotations and Accession numbers for sequences used to construct the Neighbor-Joining tree in this study.**

| Abbreviation | Annotation [organism] | Accession # |
|---|---|---|
| Cr2270(CrDhtNMT)* | 16-methoxy-2,3-dihydrotabersonine N-methyltransferase [*C. roseus*] | HM584929 |
| Cr2551* | Putative TLMT[*C. roseus*] | HM584931 |
| Cr7756* | Putative TLMT[*C. roseus*] | |
| Cr91 (CrDhtNMT) | 16-methoxy-2,3-dihydrotabersonine N-methyltransferase [*C. roseus*] | KF896244 |
| Cr706 | Putative TLMT[*C. roseus*] | KC708453 |
| Cr7756** | Putative TLMT[*C. roseus*] | KF934425 |
| Vm130 VmPiNMT | Putative TLMT[*C. roseus*] | KC708450 |
| Vm265 | Putative TLMT[*C. roseus*] | KC708451 |
| Vm2409 | Putative TLMT[*C. roseus*] | KC708452 |
| Rs820 | Putative TLMT[*C. roseus*] | KC708445 |
| Rs1755 | Putative TLMT[*C. roseus*] | KC708446 |
| Rs8692 RsPiNMT | Putative TLMT[*C. roseus*] | KC708448 |
| Rs8609 | Putative TLMT[*C. roseus*] | KC708447 |
| | | |
| Cr1196* | gamma-tocopherol methyltransferase-related protein [*C. roseus*] | HM584930 |
| HaGTMT | Gamma-tocopherol methyltransferase [*Helianthus annuus*] | ABB52798 |
| AtGTMT | Gamma-tocopherol methyltransferase [*Arabidopsis thaliana*] | AAD02882 |
| BnGTMT | Gamma-tocopherol methyltransferase [*Brassica napus*] | ACJ54674 |

*sequences found in Liscombe et al., 2010 (Note that Cr7756 is not annotated in Genbank, but the protein sequence found in the publication was used to screen for the presence of this gene in our Catharanthus databases).

**The same protein sequence as Cr7756 could be found in a root cDNA library found in Murata et al., 2006 and this has been annotated in Genbank.

**Table S3-2: Structures of MIAs used to assess substrate specificities of *CrDhtNMT*, *RsPiNMT*, and *VmPiNMT* recombinant enzymes**

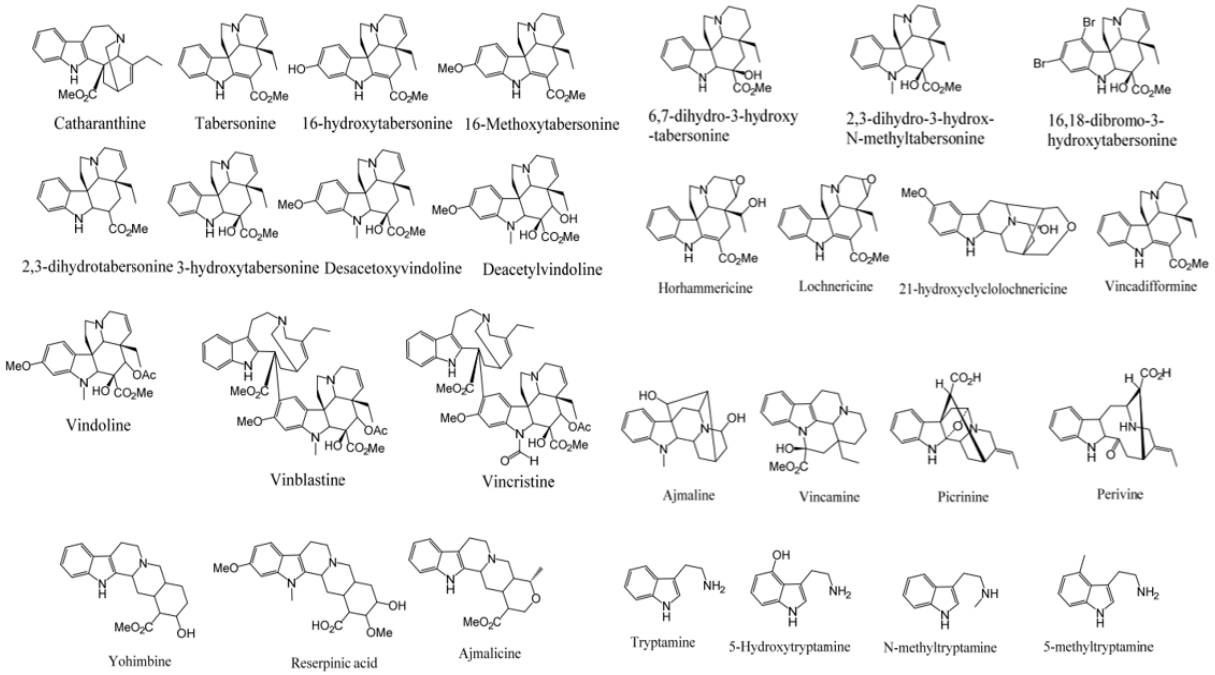


**Table S3-3: Substrate specificity analysis of *RsPiNMT* and *VmPiNMT* recombinant enzymes performed according to the standard radioactive methyltransferase assay.** *RsPiNMT* 100% activity = 0.0018 pmol product/mg protein/hour, *VmPiNMT* 100% activity = 0.045 pmol product/mg protein/hour.

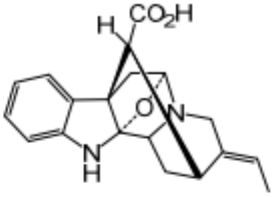
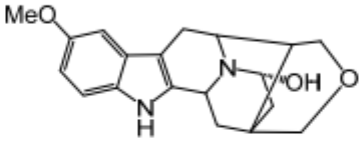
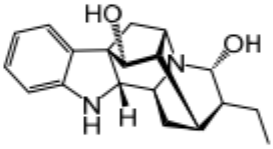
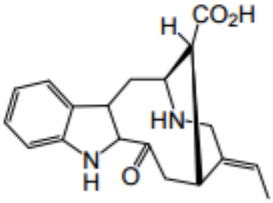
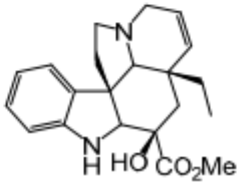
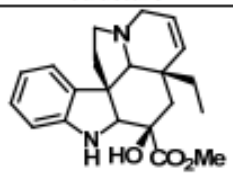
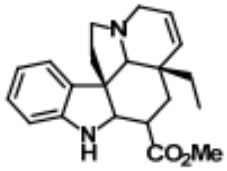
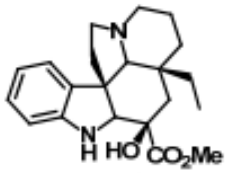
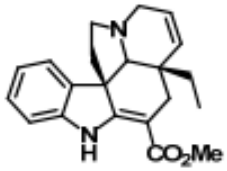
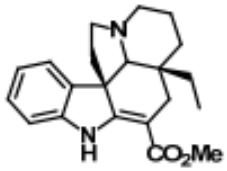
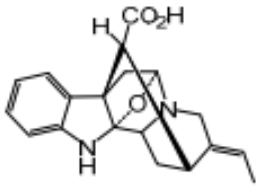
| Alkaloid | Structure | Relative activity VmPiNMT | Relative activity RsPiNMT |
|---|---|---|---|
| Picrinine |  | 100% | 100% |
| 21-hydroxylochnericine |  | 0.8% | 8% |
| norajmaline |  | 2.7% | 4% |
| perivine |  | 0% | 0% |
| 2,3-dihydro-3-OH tabersonine |  | 0% | 0% |

**Table S3-4: Substrate specificity analysis of the recombinant *CrDhtNMT* enzyme performed according to the standard radioactive methyltransferase assay.** 100% activity = 0.093 pmol product/mg protein/ hour.

| Alkaloid | Structure | Relative activity |
|---|---|---|
| 2,3-Dihydro-3-hydroxytabersonine |  | 100% |
| 2,3-Dihydrotabersonine |  | 43% |
| 2,3,6,7-Tetrahydro-3-hydroxytabersonine |  | 93% |
| Tabersonine |  | 0 |
| Vincadifformine |  | 0 |
| Picrinine |  | 0 |

# Chapter 4
# *Rauvolfia serpentina N*-methyltransferases involved in ajmaline and $N_{\beta}$-methylajmaline biosynthesis belong to a gene family derived from γ-tocopherol *C*-methyltransferase

**Authors:** Paulo Cázares Flores, Dylan Levac and Vincenzo De Luca

## 4.1 Introduction

Plant species belonging to the Apocynaceae, Loganiaceae, and Rubiaceae families contain specialized metabolic pathways responsible for the biosynthesis of a wide array of monoterpenoid indole alkaloids (MIAs) (Facchini and De Luca, 2008). MIAs are low molecular weight nitrogen-containing heterocyclic organic compounds derived from the coupling of secologanin (terpenoid moiety) and tryptamine (indole moiety). This class of plant secondary metabolites shows remarkable structural diversity and pharmaceutically valuable biological activities (O'Connor and Maresch, 2006). *Rauvolfia serpentina*, a member of the Apocynaceae family, has been extensively studied due to its ability to biosynthesize important MIAs such as ajmaline and reserpine (**Figure 4-1**). These compounds are used in the diagnosis and treatment of cardiac arrhythmias (Brugada et al., 2003) and in the treatment of high blood pressure, respectively (Anchor et al., 1955).

Pioneering studies by Joachim Stöckigt and colleagues in the early 1980's that focused on the elucidation of enzymes involved in ajmaline biosynthesis identified a metabolic pathway involving at least eight enzymatic transformations starting from the common MIA intermediate strictosidine ultimately yielding ajmaline (Wu et al., 2016). Most of these enzymes were purified to homogeneity by traditional protein chromatography from *R. serpentina* cell suspension cultures and further sequencing of oligopeptides allowed the characterization at the molecular level of their respective genes

(Ruppert et al., 2005). The remaining enzymes involved in ajmaline biosynthesis include two microsomal cytochrome P450-dependent enzymes (Smith and Stöckigt, 1995; Falkenhagen and Stöckigt, 1995), a NADPH-dependent reductase (von Schumann et al., 2002), and an *S*-Adenosyl-L-Methionine (AdoMet)-dependent *N*-methyltransferase (Stöckigt et al., 1983) that have yet to be characterized at the molecular level.

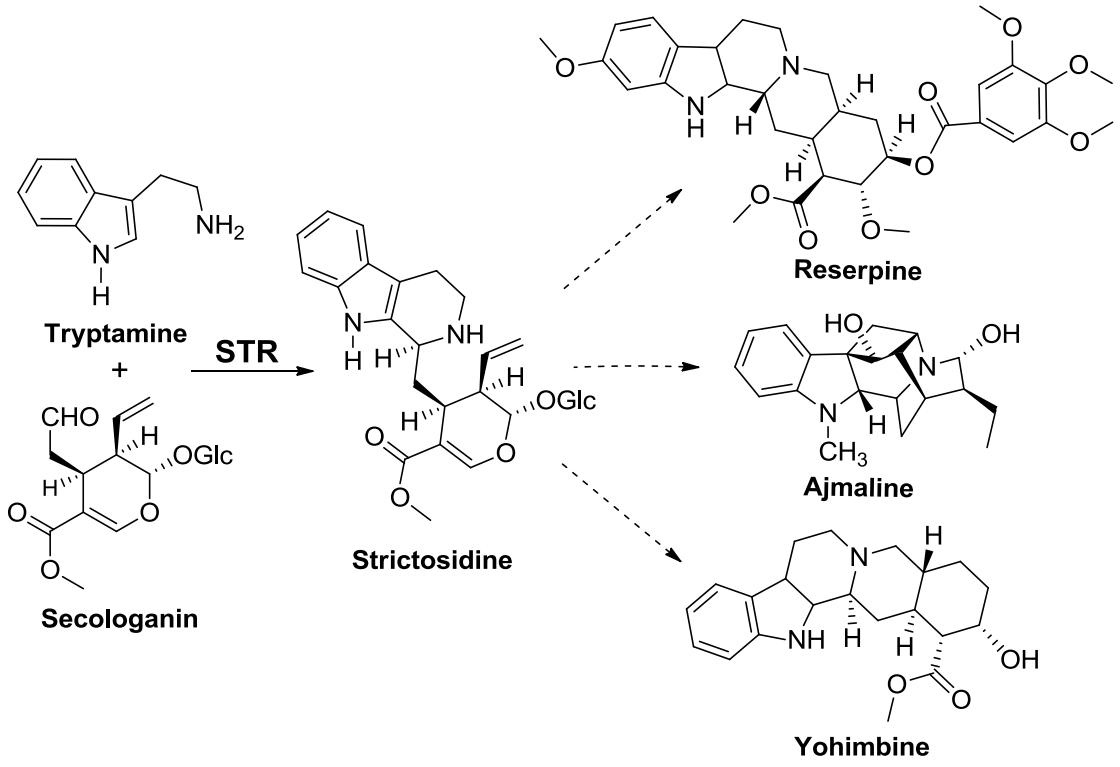



**Figure 4-1: Monoterpenoid indole alkaloids present in roots of *Rauvolfia serpentina*.** The assembly of tryptamine and secologanin to form strictosidine is mediated by strictosidine synthase (STR). Multiple enzymatic reactions (dashed arrows) mediate the formation of reserpine, ajmaline and yohimbine.

Recently, a family of AdoMet-dependent *N*-methyltransferases phylogenetically related to the γ-tocopherol *C*-methyltransferases was identified in several members of the Apocynaceae plant family (Liscombe et al., 2010; Levac et al., 2016). This previously

unknown family of γ-tocopherol-like *N*-methyltransferases (γ-TLMTs) catalyzes the *N*-methylation in MIA biosynthesis. The *C. roseus* γ-TLMT, 16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase (CrDhtNMT), is involved in vindoline biosynthesis (Liscombe et al., 2010). Analysis of evolutionary relationships of several small-molecule methyltransferases involved in specialized metabolism clustered the γ-TLMTs in a clade together with *C*-methyltransferases from vitamin E biosynthesis (Liscombe et al., 2010). The coupling of vindoline and catharanthine leads to the formation of the powerful anticancer MIA, anhydrovinblastine. The discovery of CrDhtNMT prompted the suggestion that other members of the γ-TLMT family may be responsible for the hundreds of *N*-methylated MIAs reported in the phytochemical literature (Buckingham et al., 2010). An annotated *R. serpentina* root transcript database (www.phytometasyn.ca; Facchini et al., 2012; Xiao et al., 2013) was submitted to bioinformatics analysis using the previously characterized γ-TLMT (CrDhtNMT) as a query sequence. This led to the identification of 5 γ-TLMT candidate genes that might be responsible for the final methylation step in ajmaline biosynthesis. This approach resulted in the molecular and biochemical characterization of picrinine *N*-methyltransferase as a γ-TLMT from *Rauvolfia serpentina* and a separate γ-TLMT from *Vinca minor* with the same biochemical function (Levac et al., 2016). The present study describes the molecular and biochemical characterization of two other γ-TLMT genes from *R. serpentina*, norajmaline *N*-methyltransferase (RsNNMT) that indole *N*-methylates norajmaline and ajmaline $N_{\beta}$-methyltransferase (RsANMT) that $N_{\beta}$-methylates ajmaline (**Figure 4-2**).

## 4.2 Results

### 4.2.1 Ajmaline is converted to norajmaline in biotransformations using *Streptomyces platensis* bacterial cultures

Lack of commercially available norajmaline prompted us to seek an alternative source of this MIA. Initial attempts to isolate norajmaline from several organs of *R. serpentina* such as roots, stem, leaves, and flowers failed due to low or no detectable levels observed. As a result, a microbiological *N*-demethylation of ajmaline was chosen based on the ability of *Streptomyces platensis* strain NRRL 2364 to transform commercially available ajmaline to norajmaline (Bellet and Van Thuong, 1970). Pure ajmaline standard (Extrasynthese) was fed to liquid cultures of *S. platensis* and the rate of biotransformation was monitored by ultra-performance liquid chromatography-mass spectrometry (UPLC-MS). Biotransformation experiments over ten days showed approximately 25% conversion of ajmaline (**Figure S1A**) to its *N*-demethylated form, norajmaline (**Figure S1C**). Biotransformation products were separated and purified by silica gel column chromatography to yield 95 % pure norajmaline (**Figure S4-1**) that was used together with ajmaline for biochemical characterization of *Rauvolfia* γ-TLMTs.

### 4.2.2 Recombinant proteins Rs9447 (RsNNMT) and Rs820 (RsANMT) catalyze the final steps in ajmaline and $N_{\beta}$-methylajmaline biosynthesis

The *in silico* mining of the *R. serpentina* root transcript database (www.phytometasyn.ca; Facchini et al., 2012; Xiao et al., 2013) for candidate γ-TLMTs with greater than 50% amino acid sequence identity to the CrDhtNMT (Liscombe et al., 2010) led to the selection of four candidate genes; Rs9447, Rs820, Rs8609, and Rs1755 for further analysis. The comparative amino acid analysis showed that Rs9447, Rs820, Rs8609 and

Rs1755 were 61, 54, 60 and 63 % identical to CrDhtNMT, respectively (**Figure S4-** 2 and **Figure S4-3**). Each gene was cloned and expressed in *Escherichia coli* in order to identify which of these genes was responsible for indole *N*-methylation of norajmaline. Recombinant proteins from each gene product were extracted from producing bacteria and further purified by nickel affinity chromatography (**Figure S4-4**) for enzyme assay in the presence of norajmaline and radiolabeled methyl donor *S*-adenosyl-L-(*methyl*-$^{14}$C)-methionine ([$^{14}CH_3$]-AdoMet). After separation of radiolabeled MIAs from unreacted [$^{14}CH_3$]-AdoMet, they were evaporated to dryness and dissolved in a small volume of methanol for application to silica thin layer chromatography plates. After chromatography, radioactive products were visualized using a phosphorimager. Since the Rs9447 recombinant protein converted norajmaline (Rf= 0.3) to a product that co-migrated with ajmaline (Rf= 0.5) (**Figure 4-2**, lane 1) and it did not accept ajmaline (**Figure 4-2**, lane 2) as a substrate, it was named norajmaline *N*-methyltransferase (RsNNMT). In contrast, the Rs820 protein was named ajmaline $N_\beta$-methyltransferase (RsANMT) since it converted norajmaline and ajmaline to $N_\beta$-methylnorajmaline (Rf= 0.2) (**Figure 4-2**, lane 3) and $N_\beta$-methylajmaline (Rf= 0.46) (**Figure 4-2**, lane 4), respectively. However, neither RsNNMT nor RsANMT produced a radioactive product when each was assayed in the absence of [$^{14}CH_3$]-AdoMet (**Figure 4-2**, lane 5) or MIA substrate (**Figure 4-2**, lane 6).

Additional larger scale non-radioactive assays were conducted with RsNNMT in order to produce sufficient reaction product for analysis by UPLC-MS (**Figure S4-5A-D**). Ajmaline (**Figure S4-5**; Rt= 2.3 min, m/z+327, UV max at 204, 246 and 290nm) was produced from norajmaline (**Figure S5D**; Rt=1.5min, m/z+313, UV max at 203, 237 and

285nm) in the presence of RsNNMT and AdoMet (compare **Figure S4-5A**,**B** to **C**,**D**). Indole *N*-methylation was confirmed based on similar retention time, UV absorption spectrum and mass spectrum as an authentic ajmaline standard (compare **Figure S4-5 B** to **D**). These data support the assignment of this gene as norajmaline *N*-methyltransferase.

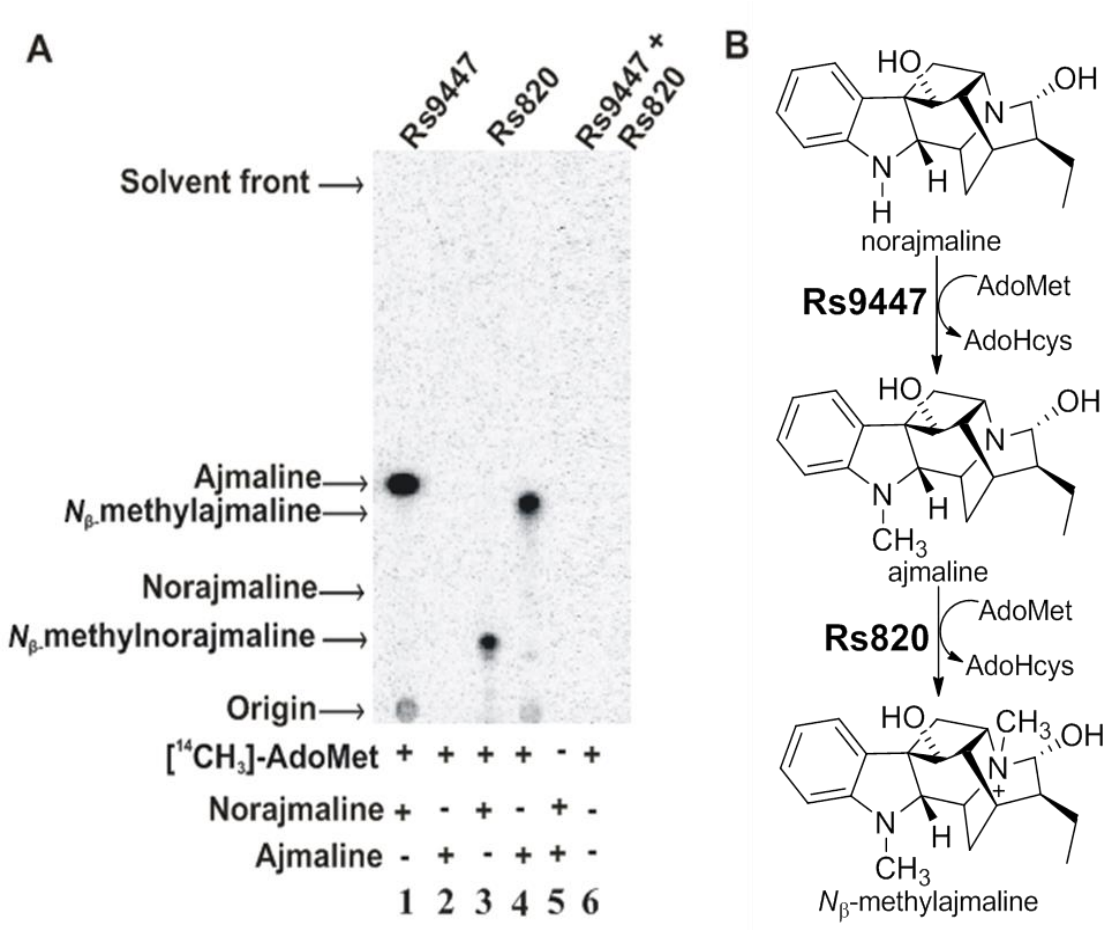



**Figure 4-2: *In vitro* characterization of the recombinant TLMT enzymes.** (**A**) Thin layer chromatography autoradiogram of the radiolabeled reaction products shows that recombinant RsNNMT (Rs9447) converts norajmaline (Rf=0.3) to a methylated product that co-migrates with an authentic ajmaline standard (Rf=0.5) (lane 1) but it does not accept ajmaline as a substrate (lane 2). Recombinant RsANMT (Rs820) accepted both norajmaline (lane 3) and ajmaline (lane 4) as substrates; however, their different Rfs ($N_\beta$-methylnorajmaline Rf=0.2 and $N_\beta$-methylajmaline Rf=0.46) suggested that transfer of a methyl group occurred in the side chain nitrogen atom (later confirmed by MS/MS shown

in **Figure S4-6**). Neither RsNNMT nor RsANMT produced a radioactive product when each was assayed in the absence of [$^{14}CH_3$]-AdoMet (lane 5) or MIA substrate (lane 6). (**B**) Proposed sequential order for the incorporation of methyl groups in biosynthesis of ajmaline and $N_\beta$-methylajmaline.

Similar larger scale non-radioactive assays were conducted with RsANMT in order to produce sufficient reaction product for analysis by UPLC-MS (**Figure S4-5 E-H**). $N_\beta$-methylajmaline (**Figure S4-5E**; Rt=2.3 min, m/z+342, UV max at 204, 244 and 287nm) was produced from ajmaline in the presence of RsANMT and AdoMet (compare **Figure S4-5E,F to G,H**). Since no standards of $N_\beta$-methylajmaline were available, this reaction product was submitted to further MS/MS analysis and the fragmentation pattern obtained revealed that the methyl group was associated to the side chain nitrogen of ajmaline resulting in a quaternary ammonium compound (compare **Figure S6 B** to **F**; m/z 110.2 to m/z 124.2), confirming the assignment of this gene as ajmaline $N_\beta$-methyltransferase.

The substrate specificities (**Table 4-1**) of recombinant RsNNMT and RsANMT showed that RsNNMT accepted norajmaline and to a lesser extent $N_\beta$-methylnorajmaline, while RsANMT accepted both norajmaline and ajmaline, with ajmaline being the best substrate for this enzyme corroborating data obtained with radioactive enzyme assays (**Figure 4-2**). RsNNMT converted norajmaline and $N_\beta$-methylnorajmaline to their respective indole *N*-methyl derivatives, while RsANMT converted norajmaline and ajmaline to their respective side chain $N_\beta$-methyl products. Neither enzyme accepted other MIA substrates tested including picrinine, perivine, 2,3-dihydrotabersonine, yohimbine or tryptamine.

Substrate specificity studies (**Table 4-1**) suggest that the best substrates for RsNNMT and RsANMT were norajmaline and ajmaline, respectively. While both enzymes were able to convert the alternate substrate to some extent as depicted in **Table 4-1**; *R. serpentina* roots appear to contain mostly ajmaline but not norajmaline, $N_{\beta}$-methylnorajmaline or $N_{\beta}$-methylajmaline. This might suggest that while the order of methylation in *R. serpentina* involves conversion of norajmaline to ajmaline and ultimately to $N_{\beta}$-methylajmaline, the terminal reaction does not occur with regularity, even if this metabolite has been reported in the phytochemical literature (Itoh et al., 2005).

To complete the partial biochemical characterization of RsNNMT and RsANMT, their pH and temperature optima were determined to be pH 7.5 and $30^{0}$C (**Figure S4-7**). Neither enzyme showed any requirement for metal ions nor was enzyme activity lost in the presence of the divalent cation chelator EDTA. Steady-state enzyme kinetic assays for recombinant RsNNMT displayed Michaelis-Menten constants ($K_m$) of 3.5 and 4.4 µM for AdoMet and norajmaline, respectively. In contrast, recombinant RsANMT had $K_m$s of 17 and 175 µM for AdoMet and ajmaline, respectively (**Table 4-2** and **Figure S4-7**). Moreover, RsNNMT displayed a higher catalytic efficiency ($k_{cat}/K_m$) for norajmaline of $477x10^3$ $s^{-1}M^{-1}$ compared to that of RsANMT for ajmaline of $5x10^3$ $s^{-1}M^{-1}$ (**Table 4-2**). The Michaelis-Menten constants and catalytic efficiencies for AdoMet and MIA substrate for RsNNMT were very similar to those observed for CrDhtNMT (Liscombe et al., 2010) and for picrinine *N*-methyltransferase (Levac et al., 2016).

**Table 4-1: Substrate specificity analysis.** Recombinant γ-TLMT enzymes were tested according to the standard radioactive methyltransferase assay in the presence of 5µg of each alkaloid substrate. RsNNMT (100% activity = 112 pkatal/mg protein); RsANMT (100% activity = 50 pkatal/mg protein).

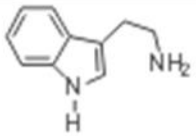
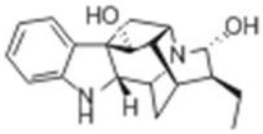
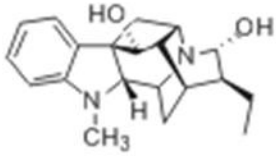
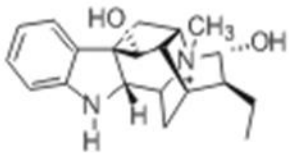
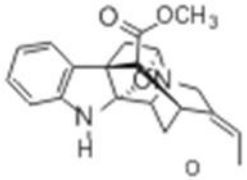
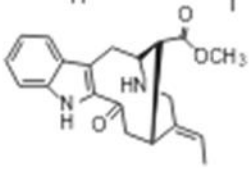
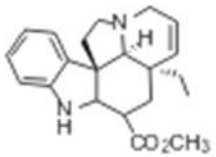
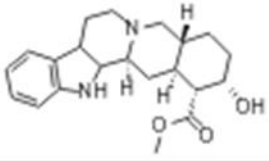
| Alkaloid | Structure | Relative activity (%) RsNNMT | RsANMT |
|---|---|---|---|
| Tryptamine (indole moiety) |  | 0 | 0 |
| Norajmaline |  | 100 | 36 |
| Ajmaline |  | 0 | 100 |
| $N_{\beta}$-methylnorajmaline |  | 18 | 0 |
| Picrinine |  | 0 | 0 |
| Perivine |  | 0 | 0 |
| 2,3-dihydrotabersonine |  | 0 | 0 |
| Yohimbine |  | 0 | 0 |

**Table 4-2: Steady-state kinetic parameters of purified recombinant γ-TLMTs.**

| Enzyme | Substrates | $K_m$ (μM) | $V_{max}$ (pmol $s^{-1}$ $mg^{-1}$ protein) | $k_{cat}$ ($s^{-1}$) | $k_{cat}$ /$K_m$ ($s^{-1}$ $M^{-1}$) |
|---|---|---|---|---|---|
| RsNNMT | AdoMet | 3.5 ± 0.2 | 96.9 ± 1.8 | 1.8 ± 0.03 | 514,286 |
| | Norajmaline | 4.4 ± 0.7 | 111.2 ± 4.7 | 2.1 ± 0.09 | 477,273 |
| RsANMT | AdoMet | 17.0 ± 2.0 | 45.7 ± 2.0 | 0.7 ± 0.03 | 41,176 |
| | Ajmaline | 184.3 ± 25.6 | 60.3 ± 3.5 | 1.0 ± 0.06 | 5,426 |

### 4.2.3 Norajmaline and ajmaline *N*-methyltransferase enzyme activities and transcripts are enriched in roots of *Rauvolfia serpentina*.

In order to study the organ specific distribution of RsNNMT and RsANMT enzyme activities, crude desalted protein extracts were obtained from different organs of *R. serpentina* and their specific activities towards their respective substrates were determined. The specific activity of RsNNMT was two to three -fold higher in roots and stems compared to other plant organs assayed (**Figure-4-3**), while RsANMT enzyme activity was detected only in roots and to a lower extent in flowers (**Figure-4-3B**). The biochemical activities correlated closely with *RsNNMT* (**Figure-4-3C**) and *RsANMT* (**Figure-4-3D**) transcript levels in different organs. However, the distribution and concentrations of ajmaline (**Figure-4-3E**) correlated more closely with the *RsNNMT* gene expression (**Figure-4-3C**) and biochemical activity (**Figure 3A**) than with those of *RsANMT* (**Figure-4-3B** and **D**). The expression patterns of RsNNMT and enzyme activity were closely correlated in *R. serpentina* leaves and were developmentally regulated with highest levels observed in the youngest leaves and decreasing progressively in older leaves (**Figure-4-3A** and **C**). The more restricted gene expression and biochemical activity of RsANMT to roots was only accompanied with mass

(m/z+342) detection of trace levels of $N_{\beta}$-methylajmaline with the appropriate retention time (Rt = 2.3 min) by UPLC-MS.

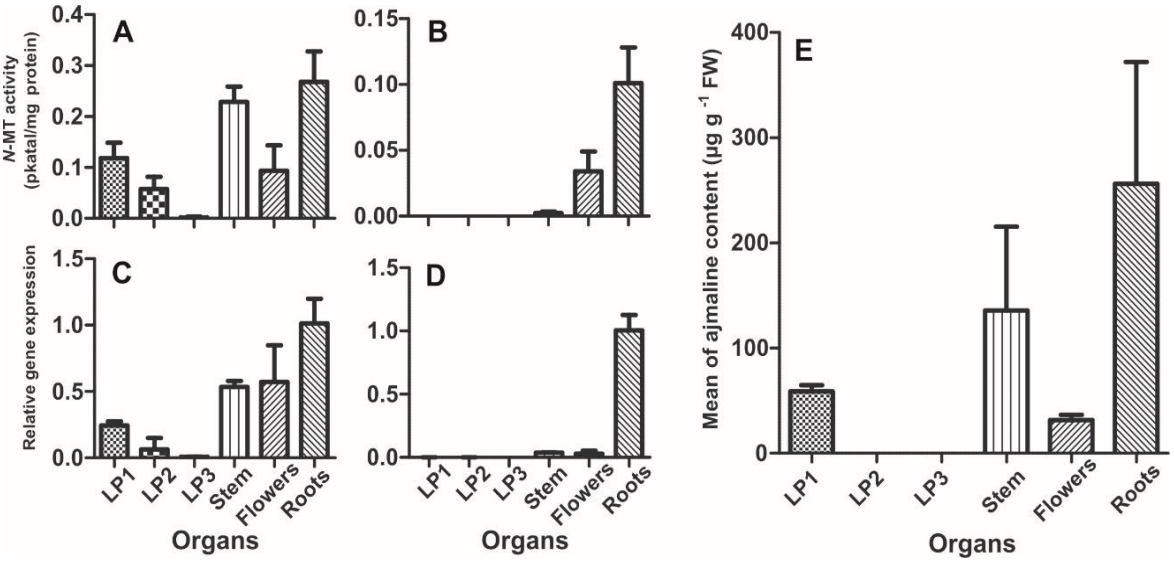



**Figure-4-3: Organ distribution of γ-TLMT enzyme activities, transcripts and ajmaline in *R. serpentina*.** Enzyme activity profile of (A) RsNNMT and (B) RsANMT in different organs of *R. serpentina*; relative transcript expression expression of (C) *RsNNMT* and (D) *RsANMT* in different organs of *R. serpentina*; (E) ajmaline content in different organs of *R. serpentina*. Error bars represent S.D. of three biological replicates. Abbreviations: LP1, youngest leaf pair next to apical meristem; LP2, older leaf pair next to LP1; LP3, older leaf pair next LP2.

### 4.3 Discussion

This study describes the homology-based molecular cloning and biochemical characterization of two *Rauvolfia serpentina* cDNAs that encode the enzymes norajmaline *N*-methyltransferase (RsNNMT) and ajmaline $N_{\beta}$-methyltransferase (RsANMT). These enzymes represent two additional members of a γ-tocopherol-like *N*-methyltransferase (γ-TLMT) gene family within the Apocynaceae plant family that are

involved in MIA *N*-methylation rather than in tocopherol biosynthesis. Members of this family display remarkable substrate specificity and high affinity for their respective MIA substrates (**Table 4-1** and **Figure S4-7**) as already described for 2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase from *Catharanthus roseus* (Liscombe et al., 2010) and picrinine *N*-methyltransferases from *Rauvolfia serpentina* and *Vinca minor* (Levac et al., 2015).

The substrate specificities for their respective MIAs and organ distribution of RsNNMT and RsANMT suggest that the indole *N* of norajmaline is methylated by RsNNMT to generate ajmaline that is commonly found within *R. serpentina* organs, while further $N_{\beta}$-methylation of ajmaline to produce $N_{\beta}$-methylajmaline may occur with much lower frequency. Although AdoMet dependent methyltransferases acting on tertiary nitrogen atoms have been previously described in protopine and benzophenanthridine alkaloid metabolism in Opium poppy (Liscombe and Facchini, 2007), ajmaline $N_{\beta}$-methyltransferase represents an unusual γ-TLMT enzyme to be characterized at the molecular and biochemical level which generates a quaternary ammonium compound in MIA biosynthesis (**Table S4-3** and **Figure S4-6**).

According to our data, we suggest that the non-detectable levels of norajmaline in organs such as roots and stems of *R. serpentina* might be due to its rapid conversion to ajmaline as a result of the high catalytic efficiency exhibited by the recombinant RsNNMT (**Table 4-2**). Furthermore, ajmaline pools were restricted to specific organs (**Figure-4-3**). Roots, stems and to some extent young leaves appeared to be the major sites of biosynthesis consistent with the organ distribution activity of the RsNNMT (**Figure-4-3**), and with previous analyses of ajmaline distribution in *R. serpentina* organs

(Roja et al., 1987; Stöckigt et a., 1997). In contrast, $N_{\beta}$-methylajmaline, a molecule structurally characterized from dried roots of *R. serpentina* (Itoh et al., 2005), was difficult to detect in our alkaloid extracts from all of the organs analyzed. Although trace amounts of $N_{\beta}$-methylajmaline in root extracts could be detected by UPLC-MS analysis, the levels were too low to determine its actual concentration. We hypothesize that low levels of this $N_{\beta}$-methylated MIA in roots may be related to the catalytic efficiency of RsANMT or to possible differences in temporal or spatial localization of the enzyme in relation to its MIA substrate. Other possible factors may have included our greenhouse conditions and developmental stage in contrast to the field-grown *R. serpentina* used in Itoh et al. (2005). Interestingly, previously characterized γ-TLMTs involved in vindoline (De Luca et at., 1987; Liscombe et al., 2010) and picrinine (Levac et al., 2016) biosynthesis were preferentially expressed in young leaves compared with RsNNMT and RsANMT that appear to be preferentially expressed in *R. serpentina* roots (**Figure-4-3**). These expression profiles clearly reflect the different organ-specific sites of biosynthesis and accumulation of different MIAs in these plant species.

In conclusion, the present study demonstrates how large-scale sequencing projects in non-model medicinal plant species are contributing to and enhancing the rate of molecular and biochemical characterization of MIA pathways in different plant species (De Luca et al., 2012; Xiao et al., 2013). The characterization of two sequential terminal reactions in ajmaline and $N_{\beta}$-methylajmaline biosynthesis was accomplished here by conducting *in silico* analysis using our annotated *R. serpentina* root's transcript database www.phytometasyn.ca (Facchini et al., 2012; Xiao et al., 2013) together with the amino acid sequence of the previously characterized γ-TLMT from *C. roseus* (Liscombe et al.,

2010) as query to pull out candidate genes. This methodology was previously also used in the characterization of γ-TLMT genes involved in picrinine *N*-methylation from *R. serpentina* and *V. minor* (Levac et al., 2016). The evolution of γ-TLMT genes from the MT family involved in tocopherol biosynthesis appears restricted to members of the Apocynaceae and may have led to the substrate-specific enzymes involved in decoration of MIAs found within this family of plants.

## 4.4 Methods

### 4.4.1 Plant material

*Rauvolfia serpentina* (L.) Benth. ex Kurz was grown in artificial soil (Sun Gro Horticulture Canada Ltd.) under controlled greenhouse conditions (22 ± $2^0$C, 70% humidity, 16:8 photoperiod) and supplemented once every month with all-purpose fertilizer [Miracle Gro, 12% nitrogen, 4% phosphate, 8% potassium (Scotts Canada)].Material from 3-month-old plants was used for developmental phytochemical, biochemical and transcriptional analysis (**Figure S4-8**).

### 4.4.2 Biotransformation (*N*-demethylation) of ajmaline by *Streptomyces platensis*

*Streptomyces platensis* (strain NRRL 2364) was purchased from the American Type Culture Collection (http://www.atcc.org). Saturated cultures of *S. platensis* were produced by growing a single colony in 10 mL of PYG media (5 g $L^{-1}$ peptone, 3 g $L^{-1}$ yeast extract and 5 g $L^{-1}$ glucose, pH 7) at $30^0$C with shaking (125 rpm) for 4-5 days. A 100 µL sample of the saturated culture was inoculated into 10 mL of Biotransformation media (3 g $L^{-1}$ $NaNO_3$, 1.3 g $L^{-1}$ $K_2HPO_4$, 0.5 g $L^{-1}$ $MgSO_4$•$7H_2O$, 0.01 g $L^{-1}$ $FeSO_4$•$7H_2O$, 0.5 g $L^{-1}$ KCl, 5 g $L^{-1}$ yeast extract, 10 g $L^{-1}$ glucose, pH 7) and incubated

under the same conditions as the initial cultures. After 5 days of growth the culture was supplemented with 10 mL of Biotransformation media and allowed to grow for one more day before adding the alkaloid substrate. Ajmaline standard (Extrasynthese) dissolved in methanol was added to the culture to a final concentration of 50 mg $L^{-1}$ and incubated under the same conditions for 8-10 additional days.

### 4.4.3 Extraction and purification of norajmaline

The final products of ten independent biotransformations (200 mL) were pooled and centrifuged at 3,000 x *g* for 5 min to pellet the cells since the media contained most of the demethylated product and remaining ajmaline starting material. After centrifugation the media was filtered through one layer of Miracloth (EMD Millipore) to ensure that no cell debris was present. Subsequently the pH of the media was adjusted to 9-10 with 5 M NaOH and extracted two times with equal volumes of ethyl acetate. The organic phase was evaporated to dryness using Rotavapor R-205 (BÜCHI) and resuspended in 500 µL of MeOH. Finally the extracted alkaloid mixture was subjected to silica gel column chromatograph using toluene:ethyl acetate:diethylamine (60:25:15 v/v/v) as a solvent system. The purity of norajmaline was confirmed by UPLC-MS (**Figure S4-1**).

### 4.4.4 Molecular cloning of *N*-methyltransferases

Using the sequencing information retrieved from our *R. serpentina* root transcripts database www.phytometasyn.ca (Facchini et al., 2012; Xiao et al., 2013), oligonucleotide primers without restriction sites were designed against the 5' and 3' UTRs of the Rs9447 gene in order to increase their specificity. For the rest of the candidate genes, primers carrying restriction sites exclusively flanked their ORFs (**Table S4-1**). cDNA generated from young *R. serpentina* roots was used as a template for PCR. The amplified gene

products were purified and cloned into the pGEM-T Easy vector (fishersci.ca) according to the manufacturer's protocol, and further subcloned into expression vector pET30-b (thermofisher.com) by restriction digest and T4-ligation procedures (Sambrock et al., 1989).

### 4.4.5 Expression and purification of the recombinant methyltransferases

A single colony of *Escherichia coli* BL21-codonPlus (DE3)-RIL (Stratagene) harboring the pET30-b/Rs9447 or pET30-b/Rs820 plasmid was grown in 100 mL of Luria-Bertani medium supplemented with 50 mg $L^{-1}$ kanamycin at 37 $^0$C with shaking (250 rpm) to an $OD_{600}$ of 0.6, and then induced with IPTG (1 mM final concentration) at 20 $^0$C with shaking (125 rpm). After ~12 h of growth at 20 $^0$C cells were collected by centrifugation (3,000 x *g*) for 15 min at 4 $^0$C. Pelleted cells were resuspended in buffer A [100 mM Tris-HCl pH (7.5), 300 mM NaCl, 10% glycerol, 14 mM β-mercaptoethanol, 10 mM imidazole] and lysed by sonication. Cell debris was removed by centrifugation (3,000 x *g*) for 25 min at 4$^0$C. Using an ÄKTA FPLC system, the supernatant was submitted to HisTrap$^{TM}$ HP (GE Healthcare) column chromatography (previously equilibrated in buffer A), which was then washed two times with 3 mL of buffer A. Proteins with low binding affinities were eluted with 3 mL of buffer A containing 100 mM of imidazole. Recombinant methyltransferases were eluted by an increasing linear gradient of imidazole (100 to 500 mM). To remove imidazole, NMT fractions were exchanged into buffer B [100 mM Tris-HCl pH (7.5), 14 mM β-mercaptoethanol] using a PD-10 Desalting column (GE Healthcare).

### 4.4.6 Steady-state enzyme kinetic analysis

Methyltransferase assays contained approximately 1 µg of purified recombinant protein, 0.48–500 µM substrate and 1.35–80 µM *S*-adenosyl-L-methionine [mixed with 0.006 µCi *S*-adenosyl-L-(*methyl*-$^{14}$C)Met 56.3 mCi mmol$^{-1}$ specific activity (perkinelmer.com)], in a total volume of 50 µL. Kinetic assays were performed under optimal conditions [30 $^{0}$C, buffer B (pH 7.5)], incubated for 15 min and terminated with the addition of 15 µL of 10 M NaOH. The reaction product was extracted with 500 µL ethyl acetate and centrifuged at 10,000 x *g* for 1 min. Finally, 250 µL of the organic phase was analyzed using a Beckman LS 6500 Scintillation counter. For substrate saturation kinetics, AdoMet was fixed at 100 µM [mixed with (0.01 µCi) *S*-adenosyl-L-(*methyl*-$^{14}$C)Met], and for AdoMet kinetics norajmaline was fixed at 100 µM and ajmaline was fixed at 500 µM. Kinetic data were fitted by non-linear regression and analyzed using GraphPad Prism version 5 for Windows (GraphPad Software). Experiments were conducted in three technical replicates.

### 4.4.7 Preparation of crude protein extracts and enzyme activity assays

Plant material from different organs of *R. serpentina* (approximately 1 g) was cut into small pieces and homogenized in 4 mL of cold buffer B using polytron (kinematica) for 10 seconds on ice. The brie was then filtered through one layer of Miracloth and centrifuged at 3,000 x *g* for 20 min at 4 $^{0}$C to remove cell debris. Supernatant was desalted using a PD-10 column chromatography. Desalted crude protein extracts (approximately 50µg) were used directly for enzyme activity assay, incubated with 10 µM norajmaline or 100 µM ajmaline, and 2.3 µM (0.02 µCi) *S*-adenosyl-L-(*methyl*-$^{14}$C)Met. Assays were performed in 150 µL reaction volume in buffer B and incubated at

30 $^0$C for 30 min. Enzyme reactions were terminated by adding 15 µL of 10 M NaOH. The product was extracted with 500 µL of ethyl acetate, and the organic phase was evaporated to dryness using SPD SpeedVac (thermoscientific.com), finally the residue was resuspended in 10 µL of MeOH containing 0.05 mg ajmaline standard and applied to TLC [Silica gel 60 $F_{254}$ (emdmillipore.com)]. TLC plates were developed in toluene:ethyl acetate:diethylamine (60:25:15 v/v/v) solvent system, and spots corresponding to ajmaline (Rf= 0.5) and $N_{\beta}$-methylajmaline (Rf= 0.46) were analyzed for radioactive incorporation in a Beckman LS 6500 Scintillation counter. Experiments were conducted in three biological replicates.

### 4.4.8 Reverse-Transcription quantitative-PCR (RT-qPCR)

Total RNA was isolated from different organs of *R. serpentina* using TRIzol® Reagent (Life Technologies) according to the manufacturer's protocol. First strand cDNA synthesis was performed using AMV Reverse Transcriptase (Promega) and oligo(dT) following the manufacturer's protocol. Gene expression experiments were performed using iTaq™ SYBR Green Supermix (Bio-Rad) under the following conditions: 95$^0$C for 3 min, then 40 cycles of 95$^0$C for 15 sec, 55$^0$C for 20 sec, 72$^0$C for 30 sec for each target gene using approximately 10 ng of cDNA and the qPCR primers listed in **Table S4-1** in a final volume of 10 µL. The expression levels were analyzed with the Bio-Rad CFX Manager Software (Bio-Rad), and normalized to the housekeeping gene [*actin* from *R. serpentina*]. To determine relative fold differences, the $C_T$ value for each target gene was normalized to the $C_T$ value for actin, and was calculated relative to a calibrator (roots) using the $2^{-\Delta\Delta C_T}$ method (Livak, K. and Schmittgen, T. 2001). Experiments were conducted in three biological replicates.

### 4.4.9 Alkaloid extraction

Plant material from *R. serpentina* (~0.5 g) was ground to a fine powder under liquid $N_2$ and incubated with 5 mL of MeOH at room temperature with wrist shaking (~100 rpm) for 30 min. The extract was centrifuged at 3,000 x *g* for 5 min to pellet cell debris. Subsequently, the supernatant was transferred to a new tube and MeOH was evaporated using the SPD SpeedVac (Thermo Savant). The residue was resuspended in 4 mL of 20% MeOH in water and subjected to acid-base alkaloid extraction performed as described by Murata, J. and De Luca, V., 2005. Finally 5 µL of each diluted sample were subjected to analysis by UPLC-MS.

### 4.4.10 Liquid Chromatography-Mass Spectrometry

Ultra-performance liquid chromatography analysis was performed using an Acquity Ultra Performance BEH C18 column with a 1.7 µm particle size, 2.1 x 50 mm dimension. The analytes were detected by photodiode array and MS. The solvent system for alkaloid analysis were as follows: solvent A [methanol:acetonitrile:5 mM ammonium acetate (6:14:80 v/v/v)] and solvent B [methanol:acetonitrile:6.2 mM ammonium acetate (25:65:10 v/v/v)]. The following linear elution gradient was used: 0-0.5 min 99% A, 1% B at 0.3 mL $min^{-1}$; 0.5-0.6 min 99% A, 1% B at 0.4 mL $min^{-1}$; 0.6-7.0 min 1% A, 99% B at 0.4 mL $min^{-1}$; 7.0-8.0 min 1% A, 99% B at 0.4 mL $min^{-1}$; 8.0-8.3 min 99% A, 1% B at 0.4 mL $min^{-1}$; 8.3-8.5 min 99% A, 1% B at 0.3 mL $min^{-1}$; and 8.5-10 min 99% A, 1% B at 0.3 mL $min^{-1}$. The mass spectrometer was operated on positive ion mode with a capillary voltage of 3.10 kV, cone voltage of 48 V, cone gas flow 2 L $h^{-1}$, desolvation gas flow 600 L $h^{-1}$, desolvation temperature of 350 $^0$C, and a source temperature of 150 $^0$C.

### 4.4.11 Tandem Mass Spectrometry (MS/MS)

All samples were run using a Bruker HCT Ultra LC/$MS^n$ system with a high capacity ion trap mass analyzer. Sample infusion was executed via a PEEK capillary from a Cole Parmer Syringe Pump and a Hamilton 250 uL syringe. Samples and standard solutions were prepared in HPLC grade Methanol and filtered through 0.2 µm PTFE syringe filters. The filtered solutions were infused by syringe pump to the nebulizer of the LC/MS system at a flow rate of 10 µL $min^{-1}$. For positive ion detection, the major parameters were set as follows: Capillary extraction voltage of 4000 V, Nebulizer (Nitrogen) 16.0 psi, Dry Gas (Nitrogen) 6.0 L $min^{-1}$ and Drying Temperature of 300 °C. For normal MS spectra, approximately 50 scans of sample data were averaged. For MS2, MS3 or MS4 spectra, usually 5-10 scans of sample data were averaged.

### 4.4.12 Protein determination:

Protein concentration of plant crude extracts and recombinant enzymes were determined as described by Bradford, M. (1976) using the protein assay dye reagent (Bio-Rad) and bovine serum albumin as standard.

## 4.5 Acknowledgements

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (V.D.L.), Canada Research Chairs (V.D.L.), Genome Canada, Genome Alberta, Genome Prairie, Genome British Columbia, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada and other government and private sector partners, P.C.F. was supported by a postgraduate PhD scholarship from Consejo Nacional de Ciencia y

Tecnología (CONACyT), México. We recognize the skilled technical work of next-generation sequencing personnel at the McGill University-Genome Québec-Innovation Centre. We are grateful to Christoph Sensen, Mei Xiao and Ye Zhang for their dedicated bioinformatics support and large scale gene annotation efforts that that helped in the identification of γ-TLMTs from the Phytometasyn web site. We also thank Tim Jones and Kyung-Hee Kim for their help in mass spectrometry analysis.

## 4.6 Supporting information

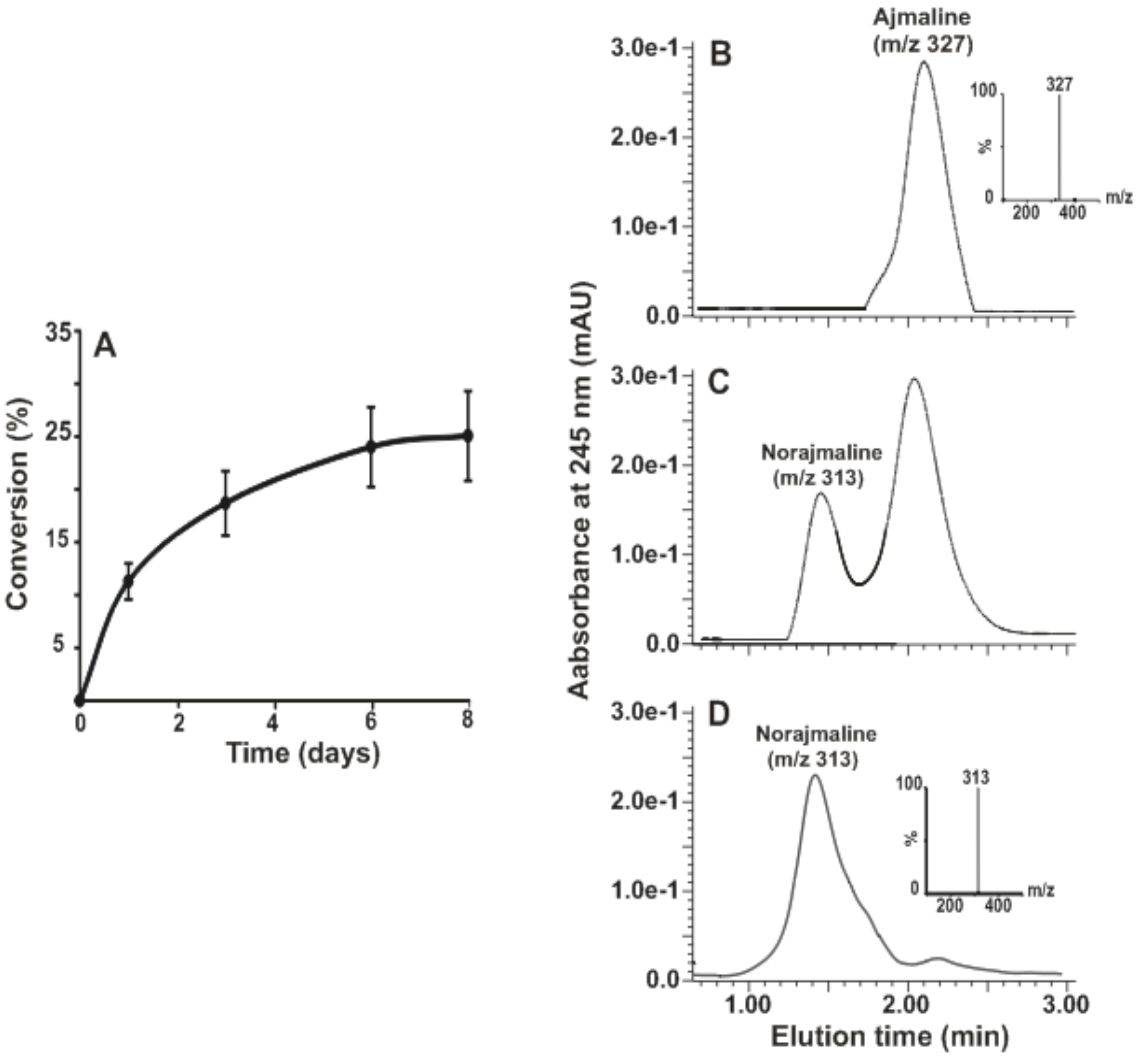



**Figure S4-1. Production of norajmaline substrate**. (**A**) Displays the conversion percentage of ajmaline to norajmaline over 8 days of biotransformation. UPLC-MS chromatograms displaying: (**B**) ajmaline standard; (**C**) *N*-demethylation of ajmaline by *S. platensis* biotransformation after 8 days; (**D**) purified norajmaline. Peak areas of the UV absorption spectrum at 245nm of an ajmaline standard curve were used to calculate the concentration of norajmaline.

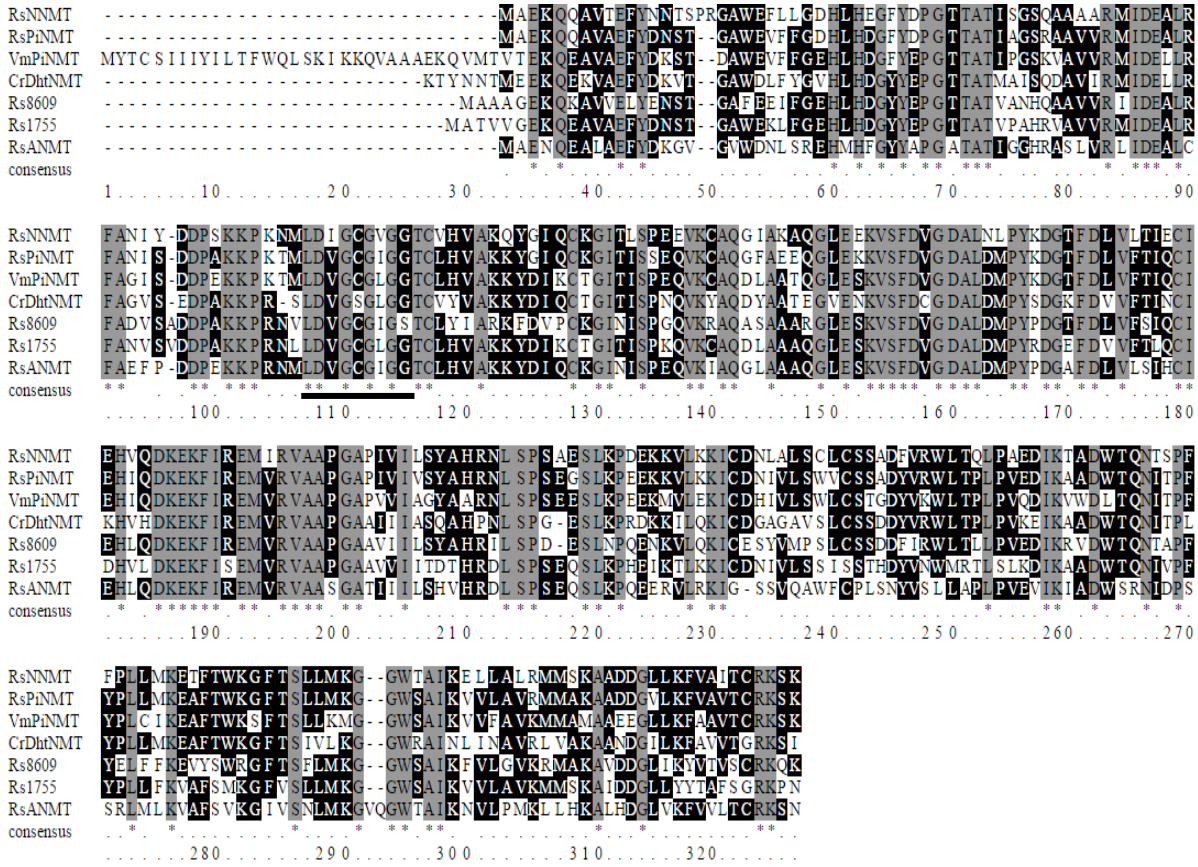



**Figure S4- 2. Multiple amino acid sequence alignment of γ-TLMT**. Alignment was performed using ClustalW 2.1(Thompson et al., 1994; http://mobyle.pasteur.fr/).Identical residues are shaded in gray background; similar residues have black backgrounds. *C. roseus* 16-methoxy 2,3-dihydro 3-hydroxytabersonine *N*-methyltransferase (CrDhtNMT) from Liscombe et al., 2012 and *R. serpentinaI* and *Vinca minor* picrinine *N*-methyltransferase (PiNMT) (Levac et al., 2016) ware included in the analysis. Motif I from MT is underline. Boxshade ClustalW analysis was performed on mobyle (Néron et al., 2009).

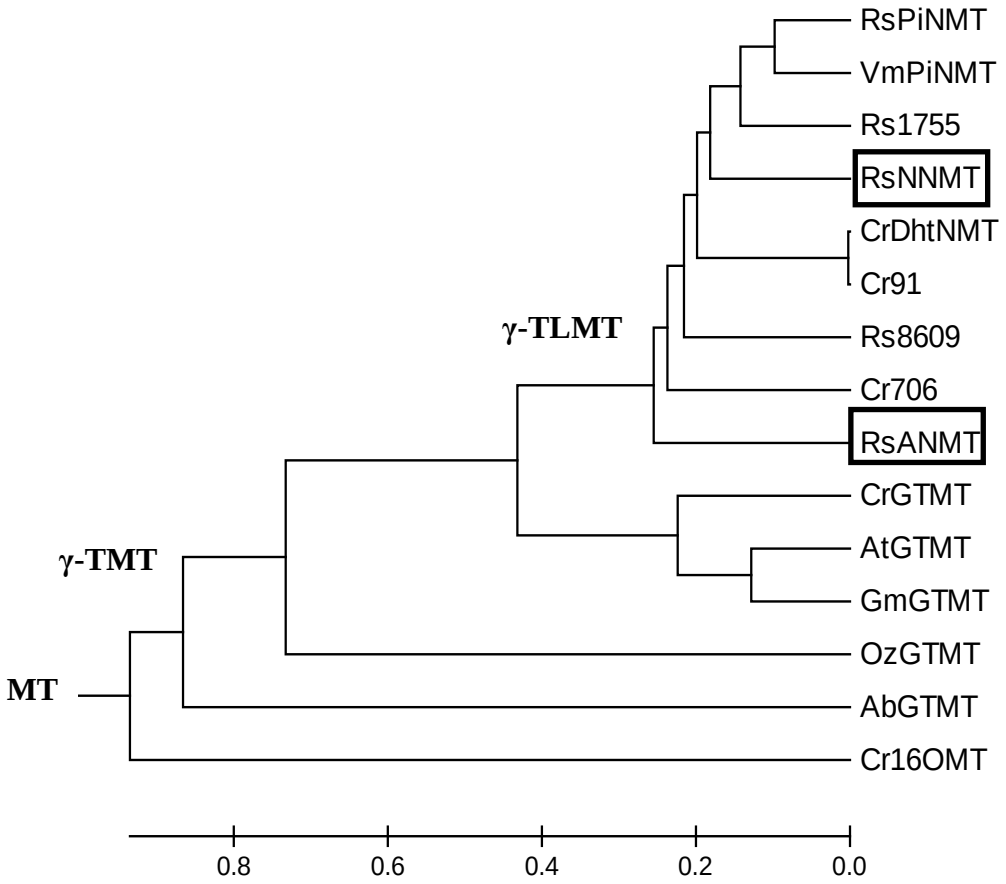



**Figure S4-3. Evolutionary relationships of γ-TLMTs.** The evolutionary history was inferred using the UPGMA method (Seneath and Sokal, 1973). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved 15 amino acid sequences (**Table S4-2**). All positions containing gaps and missing data were eliminated. There were a total of 226 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011), γ-TLMTs characterized in this study are shown in boxes.

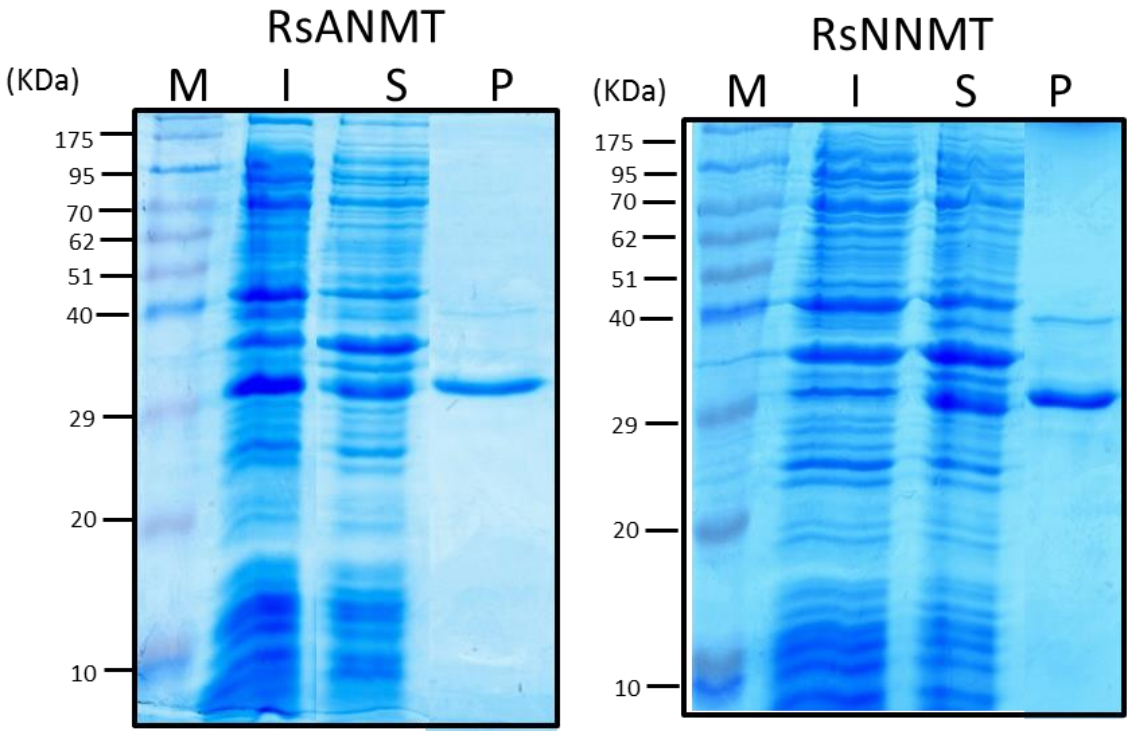



**Figure S4-4. SDS-PAGE analysis of recombinant γ-TLMT proteins.** Lanes are labeled as follows: M, marker; I, crude insoluble protein; S, crude soluble protein and P, pure protein.

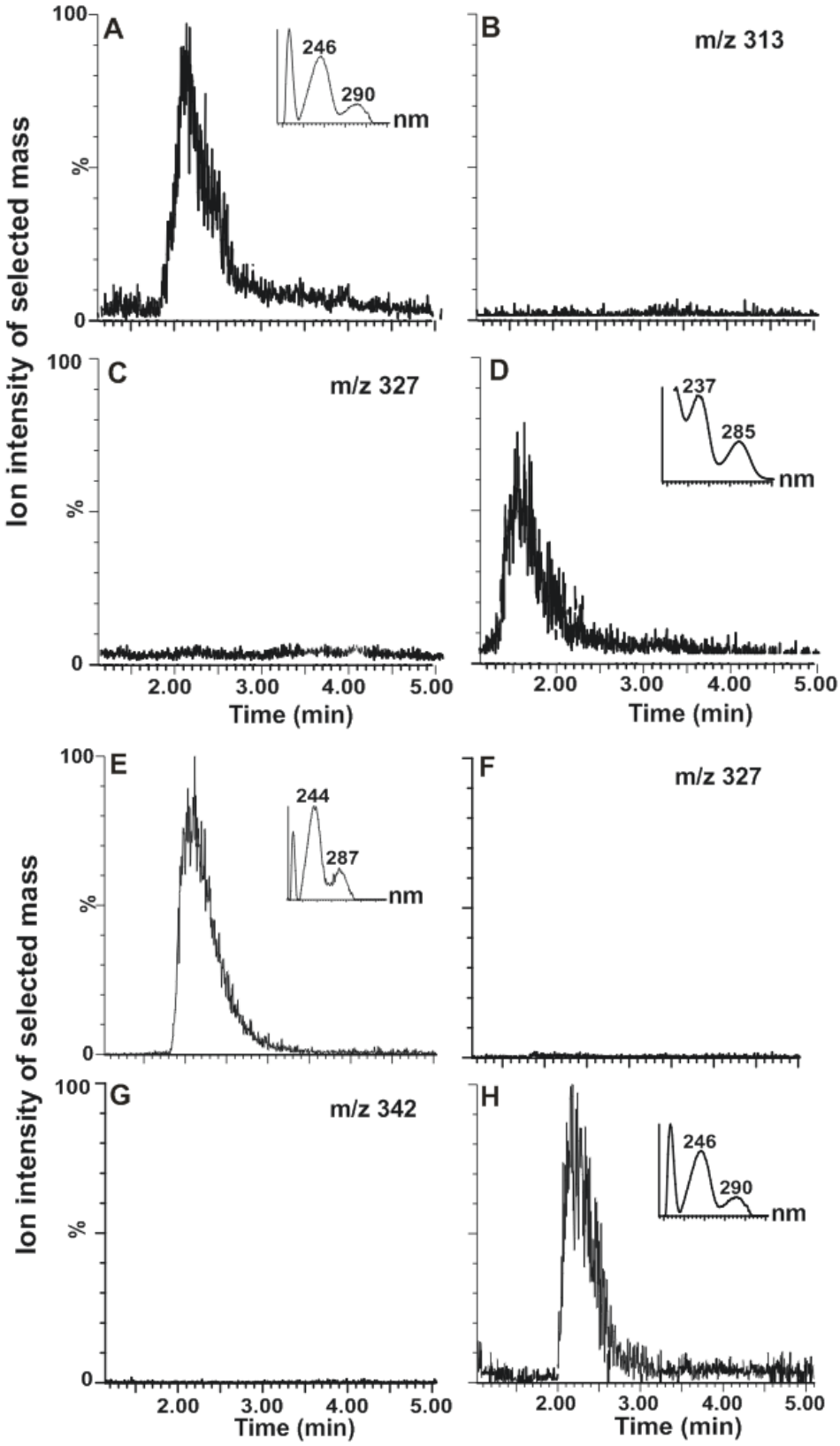

Ion intensity of selected mass
100
%
0
A
246
290
nm
B
m/z 313
C
m/z 327
D
237
285
nm
2.00
3.00
4.00
5.00
Time (min)
E
244
287
nm
F
m/z 327
G
m/z 342
H
246
290
nm

**Figure S4-5. UPLC-MS chromatograms of recombinant γ-TLMT enzyme assays.** The ion count chromatogram displays the appearance of ajmaline (A) (Rt=2.3 min, m/z+327, UV max at 204, 246 and 290nm) with the loss of norajmaline (B) (expected Rt=1.5 min, m/z+313, UV max at 203, 237 and 285nm) when RsNNMT was assayed in the presence of norajmaline and AdoMet. **Negative control**: The ion count chromatogram displays the lack of appearance of ajmaline (C) (expected Rt=2.3 min, m/z+327, UV max at 204, 246 and 290nm) with the lack of conversion of norajmaline (D) (Rt=1.5 min, m/z+313, UV max at 203, 237 and 285nm) when RsNNMT was assayed in the presence of norajmaline and absence of AdoMet. The ion count chromatogram of m/z=342, displays the appearance $N_{\beta}$-methylajmaline (E) (Rt=2.3 min, m/z+342, UV max at 204, 244 and 287nm) with the loss of ajmaline (F) (expected Rt=2.3 min, m/z+327, UV max at 204, 246 and 290nm) when RsANMT was assayed in the presence of ajmaline and AdoMet. **Negative control:** The ion count chromatogram displays the lack of appearance of $N_{\beta}$-methylajmaline (G) (expected Rt=2.3 min, m/z+342, UV max at 204, 244 and 287nm) with the lack of conversion of ajmaline (H) (Rt= 2.3 min, m/z+327, UV max at 204, 246 and 290nm) when RsANMT was assayed in the presence of ajmaline and absence of AdoMet.

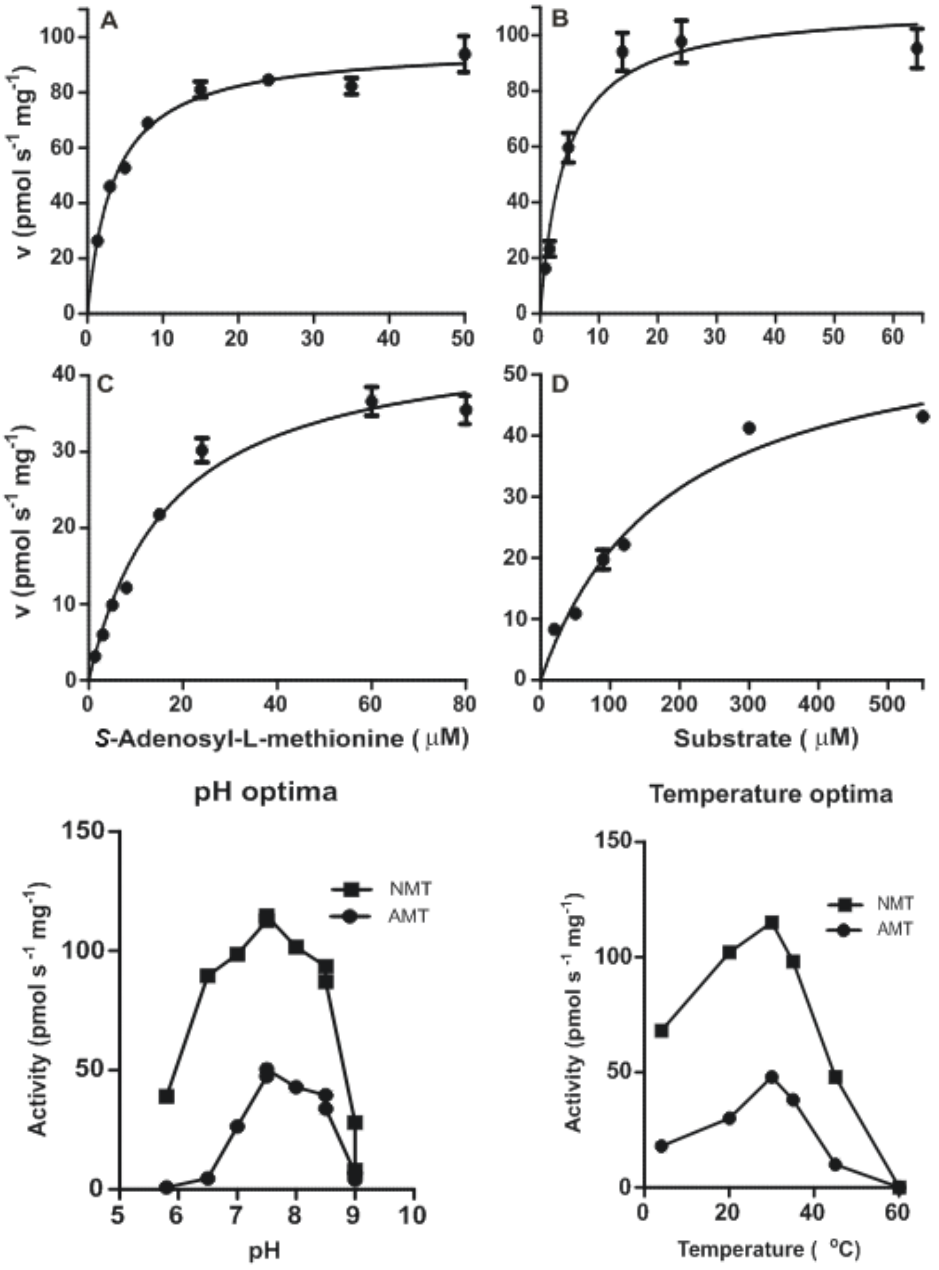



**Figure S4-6. Michaelis-Menten plots of kinetic data for purified recombinant γ-TLMTs.** (**A**) RsNNMT AdoMet saturation curve at constant norajmaline levels (100 µM). (**B**) RsNNMT norajmaline saturation curve at constant AdoMet levels (100 µM). (**C**) RsANMT AdoMet saturation curve at constant ajmaline levels (500 µM). (**D**) RsANMT ajmaline saturation curve at constant AdoMet levels (100 µM). (**E**) Curves of pH optima for RsNNMT and RsANMT were obtained from enzyme assay incubated in sodium phosphate buffer (5.8-7.5), Tris-HCl (7.5-9) and Gly-NaOH (8.5-9). (**F**) Curves

of temperature optima for RsNMT and RsANMT. Non-linear regression analysis of kinetic data was used to estimate steady-state enzyme kinetic parameters shown in **Table 4-2**.

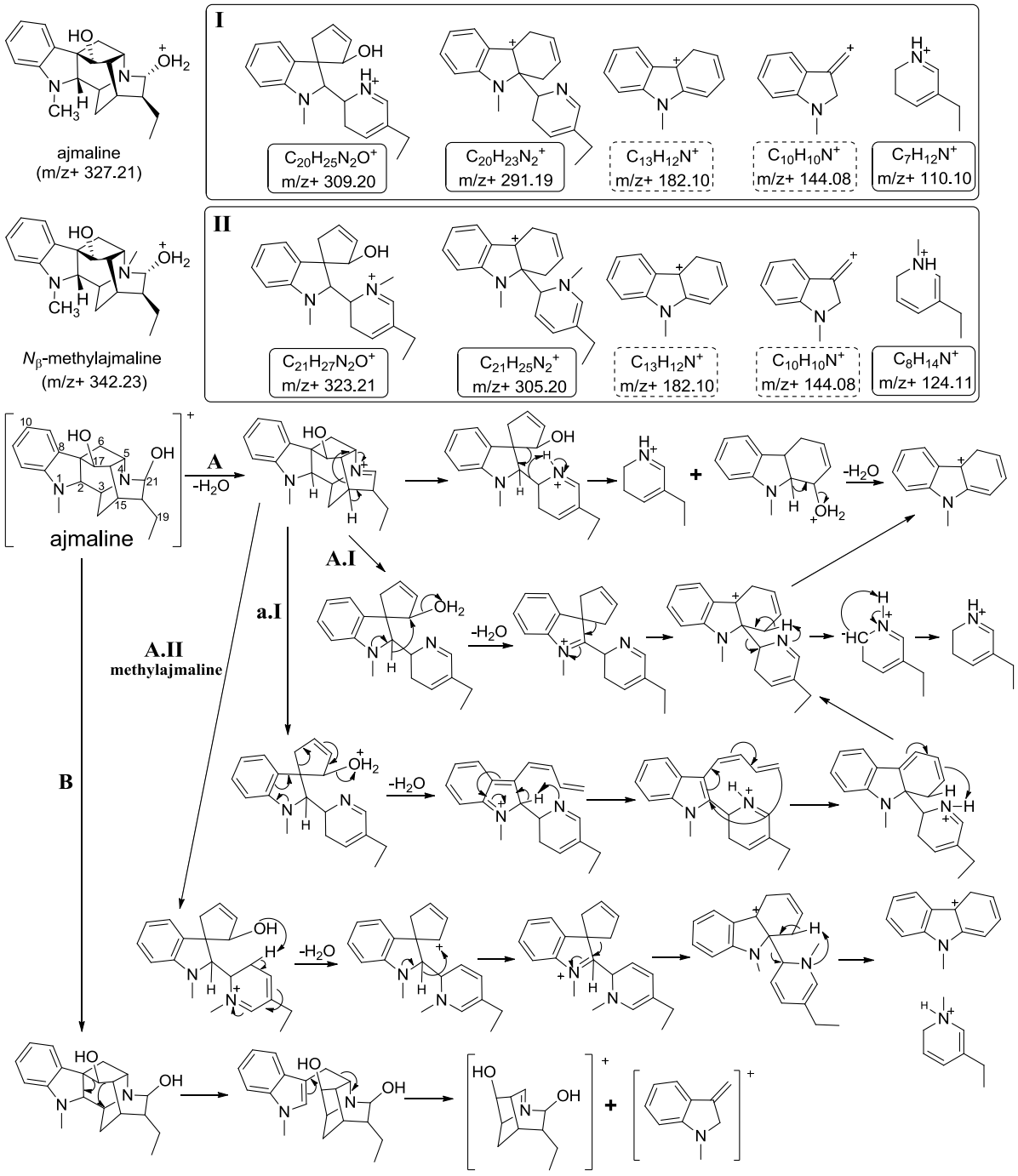
**Table S4-1. List of oligonucleotide primers used in this study.**

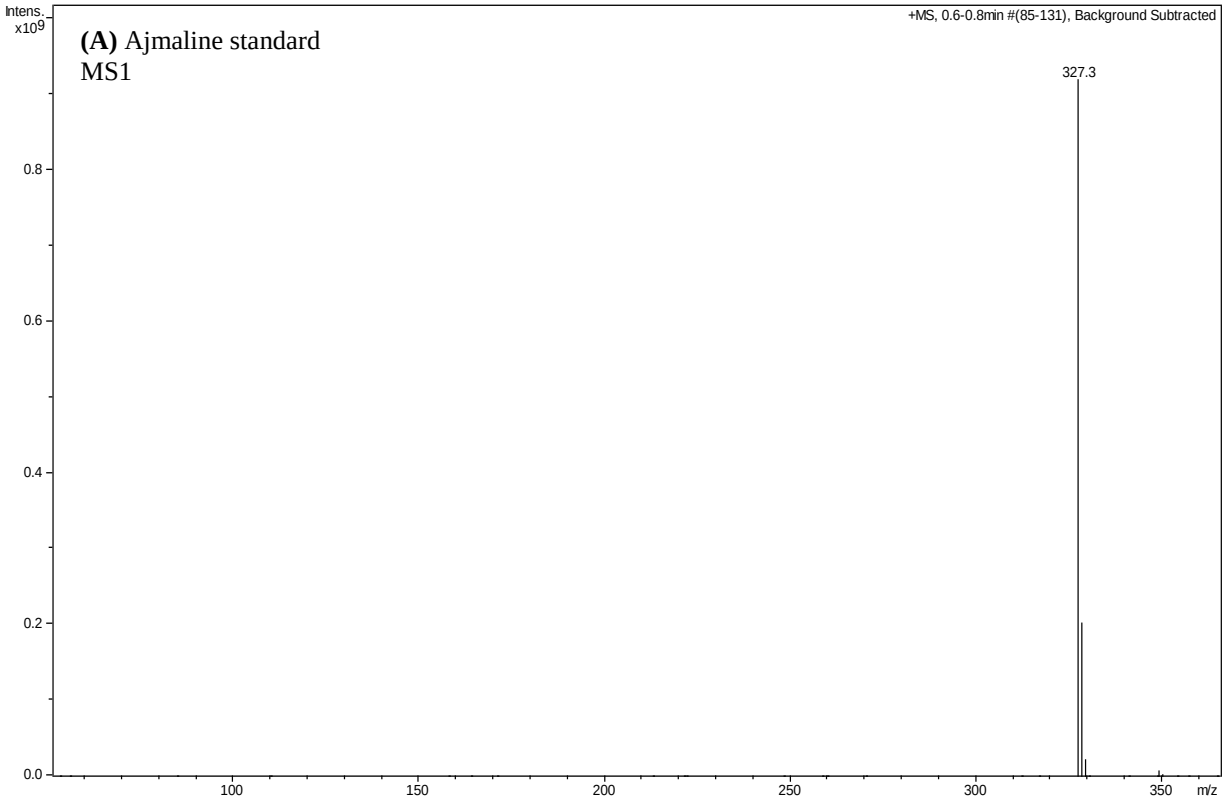
| Primer name | Sequence (5' to 3') | Purpose |
|---|---|---|
| Rs9447UTR-FW | GCAAAGTGTAAGGAGAACAACA | Full gene amplification |
| Rs9447UTR-RV | GCGATTATTTCACTACATTTATCAAGA | Full gene amplification |
| Rs9447-FW(NcoI) | TTCATATGGCAGAGAAGCAGCAGGC | Cloning |
| Rs9447-RV(NotI) | TTGCGGCCGCATTTTGATTTCCTGCATGTAATTGCAAC | Cloning |
| Rs820-FW(NcoI) | TTCCATGGCAGAGAACCAGGAGGC | Cloning |
| Rs820-RV(NotI) | TTGCGGCCGCAATTAGATTTCCGGCAAGTCAGT | Cloning |
| Rs8609-FW(NcoI) | TTCCATGGCCGCCGCTGGA | Cloning |
| Rs8609-RV(NotI) | TTGCGGCCGCATTTCTGTTTTCGACAGCTA | Cloning |
| Rs1755-FW(NdeI) | TTCATATGGCCACCGTGGTGGG | Cloning |
| Rs1755-RV(NotI) | TTGCGGCCGCAGTTAGGTTTCCGGCCG | Cloning |
| Rs9447qPCR-FW | AGATGGTTGACACAACTTCC | qPCR |
| Rs9447qPCR-RV | GCAAACTCCATAAAGCATTC | qPCR |
| Rs820qPCR-FW | TGATCCCTCTTCTCGTCTAA | qPCR |
| Rs820qPCR-RV | CCTCGTCTTACAATGAAACC | qPCR |
| Rs-actin-FW | GACCTTGAAGTACCCTATTGAG | qPCR |
| Rs-actin-RV | GTACGCCCACTAGCATATAGAG | qPCR |

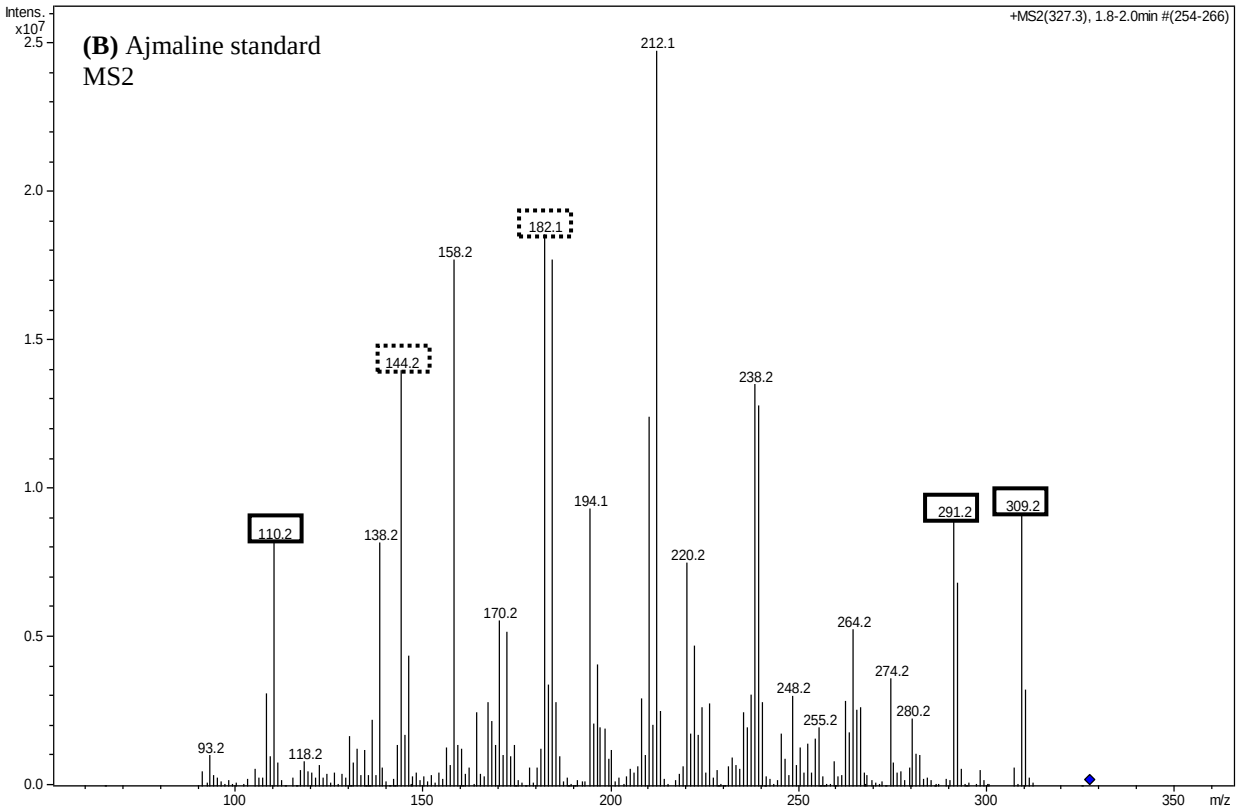
**Table S4-2. List of genes used to construct phylogenetic tree.**

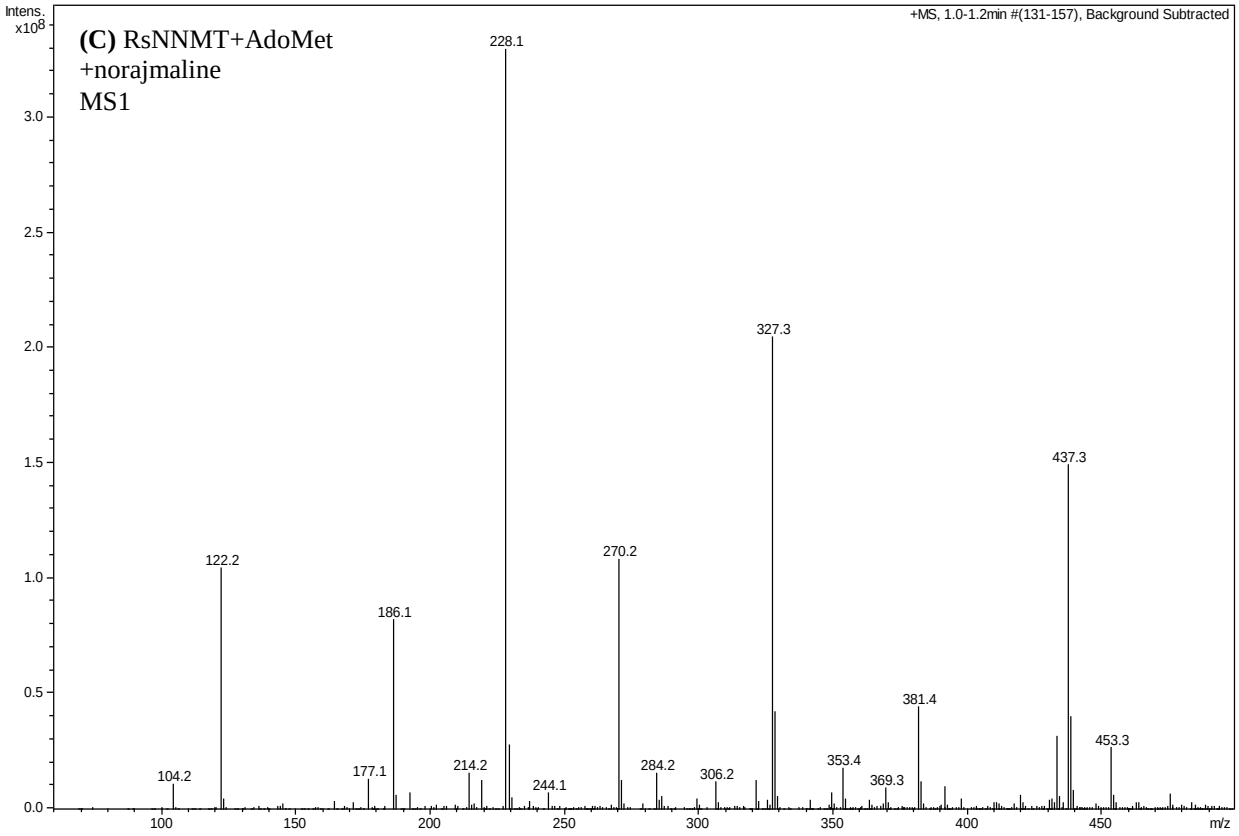
| ID | Annotation [organism] | Accession # |
|---|---|---|
| CrDhtNMT | 16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase[*Catharanthus roseus*] | ADP00410.1 |
| RsNNMT | norajmaline *N*-methyltransferase [*Rauvolfia serpentina*] | AHH02781.1 |
| RsANMT | ajmaline $N_{\beta}$-methyltransferase [*Rauvolfia serpentina*] | AHH02777.1 |
| Rs8609 | γ-tocopherol-like *N*-methyltransferase [*Rauvolfia serpentina*] | AHH02779.1 |
| Rs1755 | γ-tocopherol-like *N*-methyltransferase [*Rauvolfia serpentina*] | AHH02778.1 |
| RsPiNMT | picrinine *N*-methyltransferase [*Rauvolfia serpentina*] | AHH02780.1 |
| VmPiNMT | picrinine *N*-methyltransferase [*Vinca minor*] | AHH02782.1 |
| Cr706 | γ-tocopherol-like *N*-methyltransferase [*Catharanthus roseus*] | AHH02785.1 |
| Cr91 | 16-methoxy-2,3-dihydro-3-hydroxytabersonine *N*-methyltransferase [*Catharanthus roseus*] | AHH33092.1 |
| AtGTMT | γ-tocopherol methyltransferase [*Arabidopsis thaliana*] | NP_176677.1 |
| GmGTMT | γ-tocopherol methyltransferase [*Glycine max*] | BAK57290.1 |
| OzGTMT | γ-tocopherol methyltransferase [*Oryza sativa*] | BAF21068.1 |
| CrGTMT | γ-tocopherol methyltransferase [*Chlamydomonas reinhardtii*] | EDP02465.1 |
| AbGTMT | γ-tocopherol *C*-methyltransferase [*Acinetobacter baumannii*] | ALJ86238.1 |
| Cr16OMT | 16-hydroxytabersonine *O*-methyltransferase [*Catharanthus roseus*] | B0EXJ8.1 |

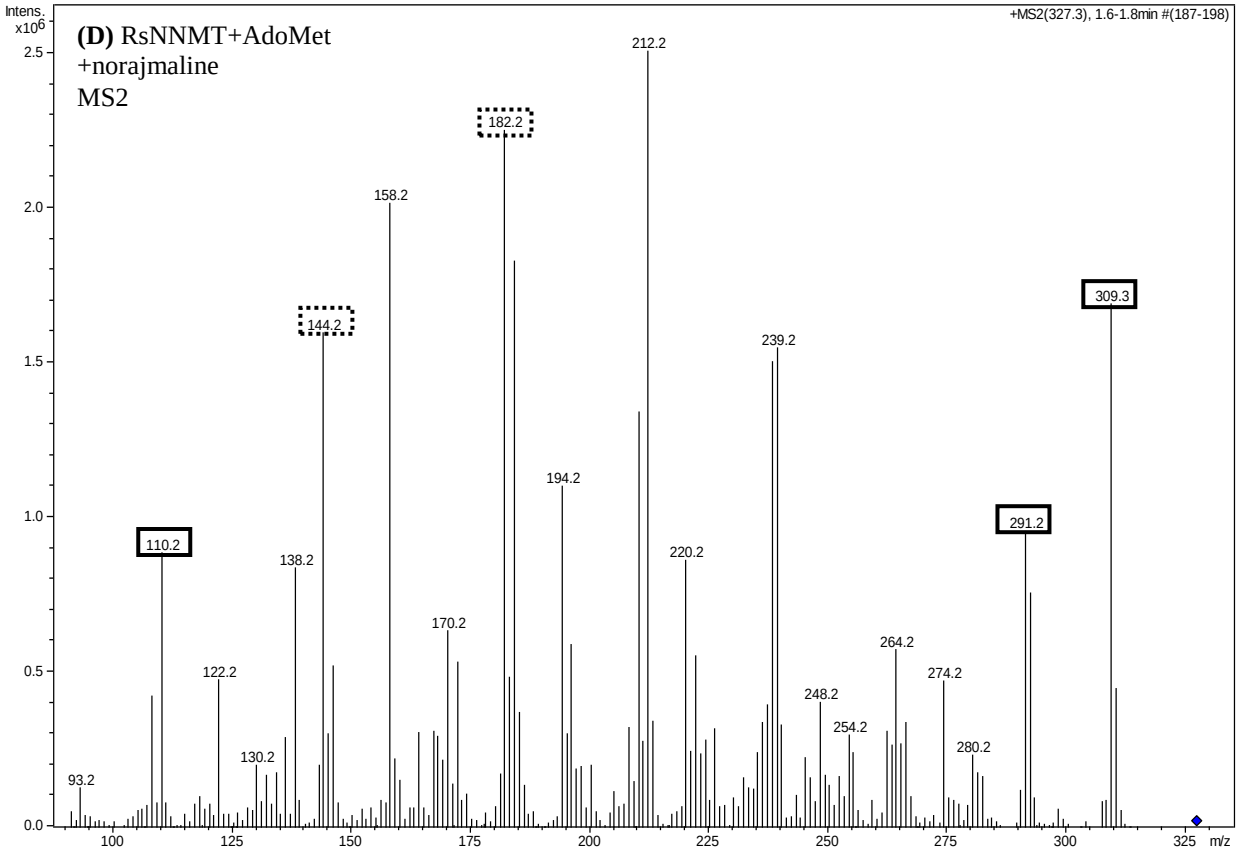
**Table S4-3. Proposed fragmentation ion pathways of ajmaline (I) and $N_\beta$-methylajmaline (II).** Exact masses shown in solid boxes were used as signatures for the identification of enzyme reaction products; masses in dashed boxes correspond to identical fragments present in ajmaline and $N_\beta$-methylajmaline.

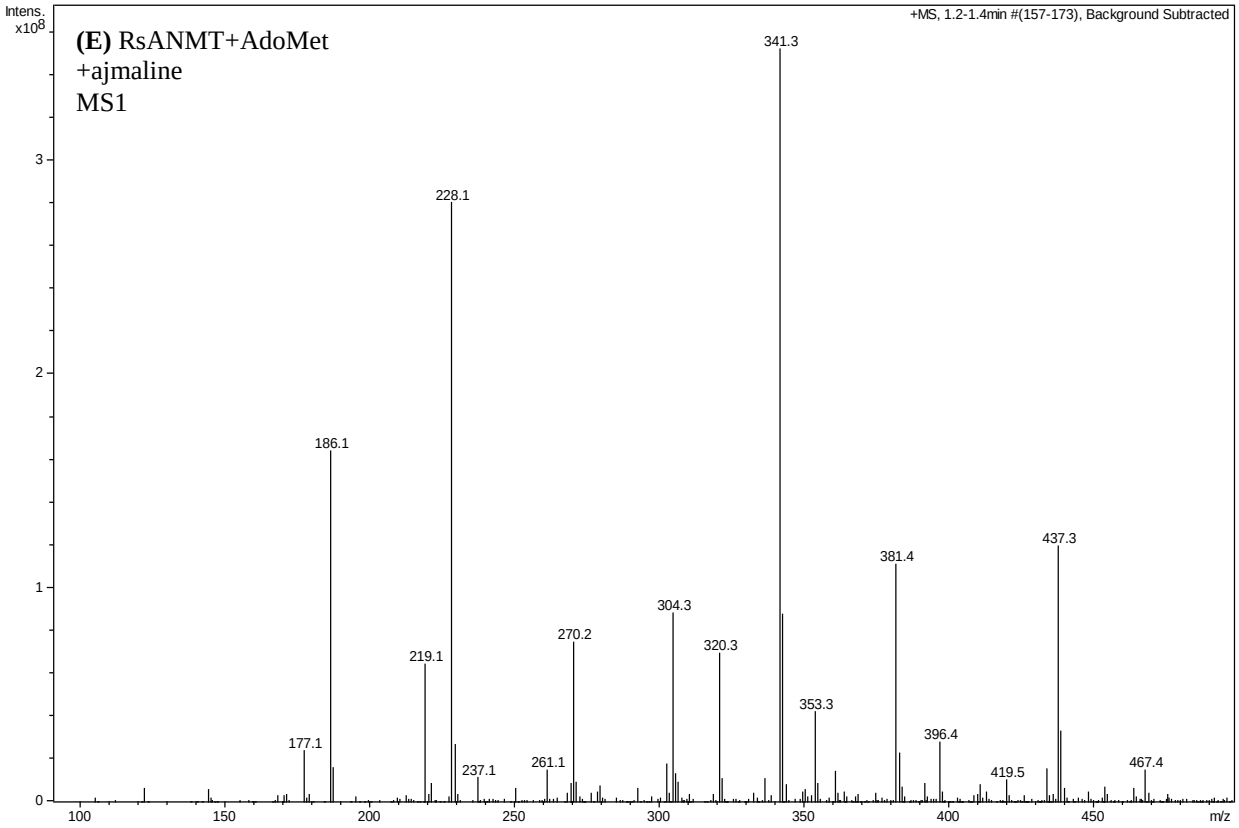



Intens.
x10^9
+MS, 0.6-0.8min #(85-131), Background Subtracted
(A) Ajmaline standard
MS1
327.3
0.8
0.6
0.4
0.2
0.0
100
150
200
250
300
350
m/z


Intens.
x10^7
+MS2(327.3), 1.8-2.0min #(254-266)
(B) Ajmaline standard
MS2
212.1
182.1
158.2
144.2
238.2
194.1
291.2
309.2
110.2
138.2
220.2
170.2
264.2
274.2
248.2
280.2
255.2
93.2
118.2
2.5
2.0
1.5
1.0
0.5
0.0
100
150
200
250
300
350
m/z


Intens.
x10^8
+MS, 1.0-1.2min #(131-157), Background Subtracted
(C) RsNNMT+AdoMet
+norajmaline
MS1
228.1
327.3
437.3
122.2
270.2
186.1
381.4
453.3
104.2
177.1
214.2
244.1
284.2
306.2
353.4
369.3
m/z


Intens.
x10^6
+MS2(327.3), 1.6-1.8min #(187-198)
(D) RsNNMT+AdoMet
+norajmaline
MS2
212.2
182.2
158.2
309.3
144.2
239.2
194.2
291.2
110.2
220.2
138.2
170.2
264.2
122.2
274.2
248.2
254.2
130.2
280.2
93.2
m/z


Intens.
x10^8
+MS, 1.2-1.4min #(157-173), Background Subtracted
(E) RsANMT+AdoMet
+ajmaline
MS1
341.3
228.1
186.1
437.3
381.4
304.3
270.2
320.3
219.1
353.3
396.4
177.1
237.1
261.1
419.5
467.4
m/z

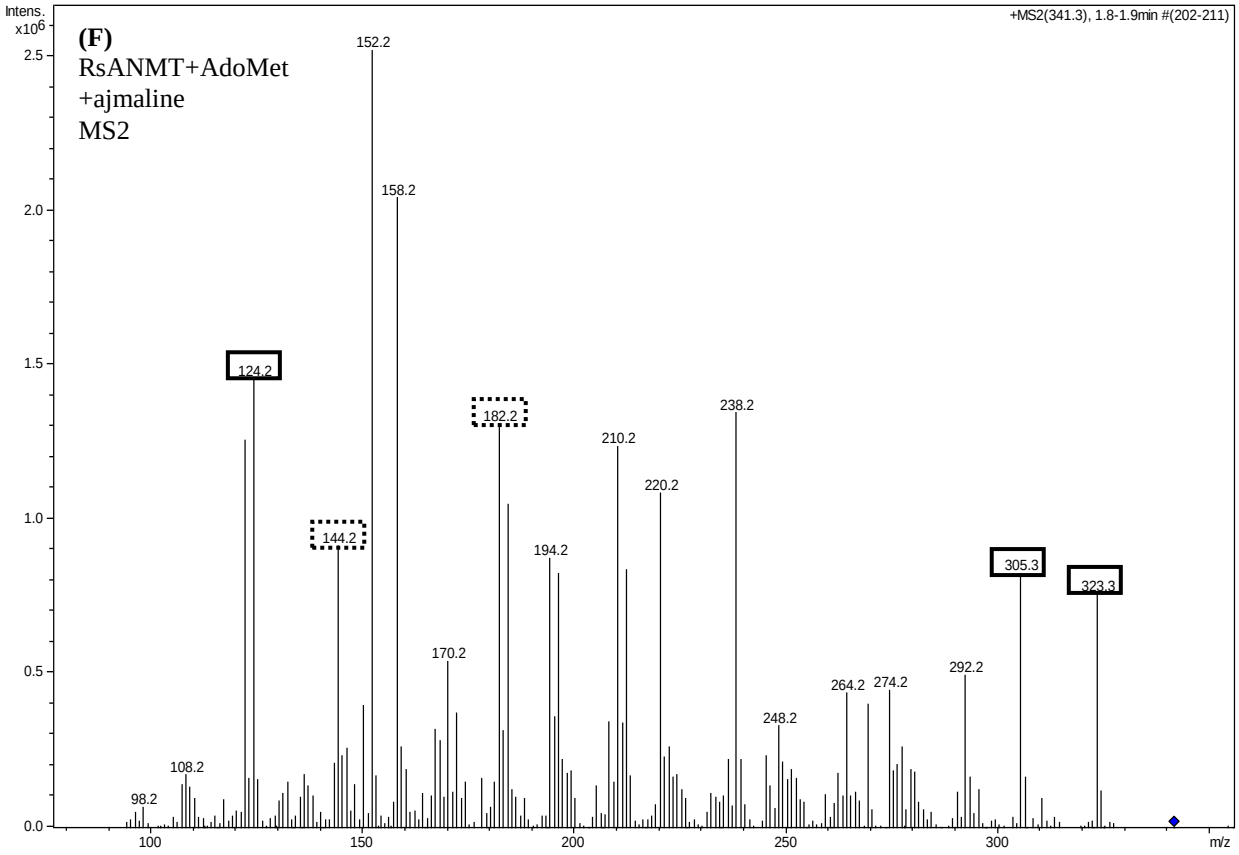



**Figure S4-7. Tandem mass spectrometry analysis.** Signature masses from **Table S3** used to assess recombinant γ-TLMT reaction products are shown in boxes. (**A**) and (**B**) MS1 and MS2 of an ajmaline standard, respectively; (**C**) and (**D**) MS1 and MS2 of ajmaline produced by RsNNMT, respectively; (**E**) and (**F**) MS1 and MS2 of $N_{\beta}$-methylajmaline produced by RsANMT, respectively.

**Figure S4-8.** ***Rauvolfia serpentina*** **plants used in this study.**

# Chapter 5
# Discussion

Monoterpenoid indole alkaloids (MIAs) are a large group of structurally diverse nitrogen-containing molecules biosynthesized by plants belonging to the Apocynaceae, Loganiaceae and Rubiaceae families. These molecules exhibit potent biological activities and have been commonly used to treat different types of illness ranging from heart disorders to cancer. Therefore, an important goal has been to elucidate the biosynthesis of MIAs in different model species at the molecular and biochemical level. In the past, much of the elucidation process relied on the ability to harvest kilograms of plant material from sources such as leaves, roots or suspended cell cultures (depending on where the target molecule was present) followed by extensive enzyme activity guided chromatographic procedures in order to obtain homogeneous enzyme preparations. However, recent developments in next-generation DNA sequencing technologies have facilitated the genome-wide interrogation of metabolic pathways, by providing access to virtually all the genes within the organism. Here we have documented the use of our large annotated transcriptome database of non-model plant species involved in the biosynthesis of MIA (*phytometasyn*) to select candidate genes through comparative bioinformatics screening followed by heterologous expression of their cDNA for functional analysis.

The discovery of a *Catharanthus roseus S*-Adenosyl-L-Methionine-dependent enzyme catalyzing the indole *N*-methylation of 16-methoxy 2,3-dihydro 3-hydroxytabersonine involved in the biosynthesis of vindoline, an important constituent of vinblastine and vincristine, MIAs used in leukemia treatment, provided our query sequence to mine the transcriptomic databases of plant species specialized in the synthesis of *N*-methylated MIAs. This approach was further validated by the molecular

cloning and functional expression of two separate picrinine *N*-methyltransferases from *Vinca minor* and *Rauvolfia serpentina*, respectively. In order to identify the potential substrates for these recombinant enzymes, we used a targeted enzyme-activity profile approach. Soluble recombinant enzyme isolates were incubated with a mixture of alkaloids extracted from the leaf surfaces of MIA-producing plants and incubated with a radiolabeled cofactor (*S*-Adenosyl-L-[*methyl*-$^{14}$C]-Methionine). This approach allowed us to track the incorporation of the methyl group through thin layer chromatography followed by autoradiography to visualize radioactive spots. This approach led to identification of picrinine as a substrate for both recombinant enzymes. Recombinant enzymes were able to catalyze the indole *N*-methylation of picrinine converting it to ervincine (*N*-methyl-picrinine). We confirmed that transfer of the methyl group occurred at the indole nitrogen moiety by structural analysis of the purified reaction product using ultraviolet (UV) and infrared (IR) spectroscopy. Furthermore, we determined the enzyme kinetic parameters of the purified recombinant proteins and tested their substrate specificity. Our results showed that *R. serpentina* and *V. minor* picrinine *N*-methyltransferase recombinant proteins have high affinity for picrinine and were highly selective for substrates bearing a cyclic ether ring system. Moreover, molecular phylogenetic analysis clustered these enzymes in a monophyletic clade separated from closely related γ-tocopherol *C*-methyltransferases, thus, the assignment as a new family of γ-tocopherol-like *N*-methyltransferases (γ-TLMT) involved in nitrogen methylation of MIA.

Additionally, we cloned two more members of the γ-TLMT family from *R. serpentina* in order to identify the unknown *N*-methyltransferase genes involved in

ajmaline biosynthesis, by following the same approach used for the identification of picrinine *N*-methyltransferase. Interestingly, we found that one of the *R. serpentina* recombinant enzymes was able to methylate the side chain nitrogen atom of ajmaline in order to produce $N_{\beta}$-methylajmaline carrying a positively charged quaternary nitrogen. This molecule represents an unusual MIA found in the roots of *R. serpentina*. We confirmed that methylation takes place on the side chain nitrogen rather than the indole nitrogen by producing titer amounts of the enzymatic reaction product and subjected it to further tandem mass spectrometry analysis (MS-MS). In functionally characterizing the enzyme responsible for the last step in ajmaline biosynthesis, which involves the indoline *N*-methylation of norajmaline to produce ajmaline, we faced a different challenge: substrate availability. Initial search for the norajmaline substrate across different organs of *R. serpentina* did not produce any measurable quantities. Therefore, we decided to explore a microbiological *N*-demethylation reaction based on the ability of *Streptomyces platensis* to transform ajmaline into norajmaline. The *S. platensis* biotransformations provided us with sufficient amounts of norajmaline to carry out further biochemical assays with the γ-TLMT recombinant enzyme candidates. We found that indoline *N*-methylation of norajmaline was carried out by one of our γ-TLMT clones, making a stronger case for the previous assumption that members of this newly discovered γ-TLMT family are responsible for the hundreds of *N*-methylated MIA found in nature. Additionally, substrate specificity assays revealed that indoline *N*-methylation of norajmaline takes place first, followed by side chain *N*-methylation of ajmaline, delineating the enzymatically favorable order of the reactions. Furthermore, we provided a comprehensive overview of all nine functionally characterized biosynthetic gene

transcripts across organs and leaf developmental stages in *R. serpentina*, revealing that ajmaline biosynthesis is subjected to developmental regulation in leaves. The studies documented here illustrate the use of bioinformatics guided screen to identify candidate genes involved in specialized metabolism, followed by targeted enzyme-activity profile to assess their biochemical function. Moreover, we provide an interesting approach to access substrates through microbial biotransformations.

# Chapter 6
# Conclusions and Future Work

We have shown how large scale DNA-sequencing projects in addition to bioinformatics-guided cloning and targeted enzyme-activity profile are enhancing and contributing to the molecular and biochemical characterization of enzymes involved in specialized metabolism. Our studies provide additional biochemical and molecular information of a recently identified group of enzymes responsible for the indole and side chain *N*-methylation of monoterpenoid indole alkaloids within the Apocynaceae, the γ-tocopherol-like *N*-methyltransferase (γ-TLMT) family. We have determined that the last two enzymatic steps involved in the ajmaline metabolic pathway are carried out by members of the γ-TLMT family, closing a gap in our knowledge of the terminal reactions of this important biosynthetic pathway in *R. serpentina*. Future work on *R. serpentina* will focus on the identification of the remaining four unknown biosynthetic enzymes involved in ajmaline biosynthesis by using similar approaches to the ones documented in this thesis. The two candidate NADPH-dependent reductase genes: vomilenine and 1,2-dihydrovomilenine reductases have already been cloned into a plasmid expression vector pET30b by Dylan Levac, and I have optimized the expression conditions for bacterial cultures producing recombinant enzymes. Further experiments will involve the isolation and purification of vomilenine substrate from *R. serpentina* organs, such as roots and leaves, for enzyme activity assays and biochemical characterization. In functionally characterizing the cytochrome P450-dependent enzymes: sarpagan bridge enzyme and vinorine hydroxylase, we propose a comparative bioinformatics approach for the selection of candidate genes followed by their molecular cloning and targeted enzyme-activity profile.

# Appendix A
# Ajmaline biosynthesis is developmentally regulated in *Rauvolfia serpentina* leaves

**Authors:** Paulo E. Cázares Flores and Vincenzo De Luca

## A.1 Introduction

The biosynthetic steps leading to the formation of the antiarrhythmic monoterpenoid indole alkaloid (MIA) ajmaline have been the subject of study for over three decades (Ruppert et al., 2005a). Detailed mechanistic analyses in *Rauvolfia serpentina* cell suspension cultures have resolved a metabolic pathway involving at least eight enzymatic transformations (**Figure 2-3**). Biochemical assays suggest that a biogenetic intermediate towards the biosynthesis of sarpagan and ajmalan MIA-backbones could be 4,21-dehydrogeissoschizine or 4,5-dehydrogeissoschizine (Schmidt and Stöckigt, 1995; Ingham et al., 2012; O'Connor and Maresh, 2006). This intermediate is produced immediately after removal of the glucose moiety from strictosidine, mediated by strictosidine β-glucosidase (**Figure 2-3**). Although no gene has yet been assigned for the sarpagan-bridge forming enzyme, its biochemical activity was exclusively detected in microsomal preparations from *R. serpentina* cell cultures, revealing the involvement of a membrane bound cytochrome P450-dependant monooxygenase to generate polyneuridine aldehyde carrying the sarpagan-bridge between C-5 and C-16 in MIA biosynthesis (Schmidt and Stöckigt, 1995). Following the formation of polyneuridine aldehyde, polyneuridine aldehyde esterase (Pfitzner and Stöckigt, 1983) mediates methyl group hydrolysis to form a reactive carboxylic acid that undergoes spontaneous decarboxylation and generates an aldehyde group attached to C-16 forming 16-epi-vellosimine (**Figure 2-3**). Importantly, this reaction represents the first committed step towards the formation

of sarpagan and ajmalan MIA backbones. Depending on the stereo configuration of the aldehyde group, 16-epivellosimine directs the formation of ajmaline while vellosime directs the formation of sarpagine (**Figure 2-3**) (Pfitzner and Stöckigt, 1983). The next step in the ajmaline pathway is carried out by vinorine synthase, an acetyltransferase generating the ajmalan ring between C-7 and C-17 (Pfitzner et al., 1986; Ma et al., 2005). This enzyme accepts 16-epivellosimine exclusively to form vinorine containing the six-membered rings typical of all ajmalan-type MIAs (**Figure 2-3**). Vinorine is then hydroxylated at C-21 by an unidentified cytochrome P450-dependent monooxygenase (Falkenhagen and Stöckigt, 1995) to generate vomilenine. Saturation of the 1,2-indolenine and 19,20 double bonds of vomilenine represent the next biosynthetic steps in the ajmaline pathway. These reactions are mediated by two separate NADPH-dependent reductases (von Schumann et al., 2002; Gao et al., 2002). Peptide sequencing of the purified reductases has afforded identification of their respective cDNAs (Ruppert et al., 2005a). Nevertheless, recombinant expression of these enzymes for further biochemical analysis still needs to be addressed. The saturated product, 17-*O*-acetylnorajmaline, is then deacetylated by acetylajmalan esterase in order to generate norajmaline (Polz et al., 1987; Ruppert et al., 2005b). The final step involves the indoline nitrogen methylation of norajmaline to produce ajmaline. This reaction is carried out by an *S*-Adenosyl-L-Methionine-dependent *N*-methyltransferase member of the γ-tocopherol-like *N*-methyltransferase (γ-TLMT) family (see **Chapter 4**).

To date, the ajmaline metabolic pathway has been exclusively investigated in cell suspension cultures of *R. serpentina*. In order to provide a complete understanding of this important metabolic pathway, we examined the relative transcript abundance of the

functionally characterized genes involved in ajmaline biosynthesis by performing reverse-transcription quantitative polymerase chain reaction (RT-qPCR) analysis in different organs and developmental stages of *R. serpentina* leaves. We used our large annotated *R. serpentina* roots transcriptome assembly database (http://www.phytometasyn.com/; Facchini et al., 2012; Xiao et al., 2013) for molecular analysis of the cDNAs encoding the characterized ajmaline metabolic enzymes. Furthermore, we validated our RT-qPCR data by analyzing the phytochemical constituents in each *R. serpentina* organ. Our results indicated that transcripts of the biosynthetic genes involved in ajmaline biosynthesis are subjected to developmental regulation in leaves, activating in early leaf developmental stages and decreasing progressively as the leaf ages.

## A.2 Results

### A.2.1 Molecular analysis of genes involved in ajmaline biosynthesis – qPCR primer design and validation

In order to design qPCR oligonucleotide primers for each target biosynthetic gene, we pulled out the cDNA sequences encoding nine steps in ajmaline biosynthesis (**Table A-1**) from our annotated *R. serpentina* large transcriptome assembly database (www.phytometasyn.com/). All of these sequences represented complete transcripts including untranslated regions (UTRs). Subsequently, we copied the transcript sequences into the software A plasmid Editor (ApE: biologylabs.utah.edu/jorgensen/wayned/ape/) and used the find-primer-tool to design the appropriate qPCR primers. Moreover, to increase the specificity of our qPCR primers we decided to flank sequences that contained part of a UTR as well as the open reading frame (ORF). To validate the

designed qPCR primers, we performed regular PCR using cDNA generated from *R. serpentina* roots as template. All of these reactions gave positive products with the expected length (**Figure A-1**).

**Table A-1: List of oligonucleotide primers used to determine transcript abundance of ajmaline biosynthetic genes in *R. serpentina.***

| ID | Description [Genbank accession #] | qPCR primers (5'-3') | Product length (bp) |
|---|---|---|---|
| *TDC* | Tryptophan decarboxylase [N/A] | FW- CTTTCTGCTTCATACACCTCC<br>RV- ACAGGATAGCTCTCGACATTC | 222 |
| *STR* | Strictosidine synthase [X62334.1] | FW- TTAAATCCCAGTCCTTCGTC<br>RV- AGTTGGGTCCTTCGTACTTG | 250 |
| *SGD* | Strictosidine β-glucosidase [AJ302044.1] | FW- TGTATGGTCATTCTTCGACAAC<br>RV- GCTTATTGAAATGCCCAAAC | 267 |
| *PNAE* | Polyneuridine aldehyde esterase [AF178576.1] | FW- AAAAGTTCTCAACCGAGAGG<br>RV- CTCACATTGTTTGCTCCCTTC | 252 |
| *VS* | Vinorine synthase [AJ556780.2] | FW- CCAATGGCAGAAGATGAAATG<br>RV- ATACATCCTAGTGCGGTAGTG | 206 |
| *VR*[a] | Vomilenine reductase [N/A] | FW- ATGTAAACACTGCAATGGAGC<br>RV- TGAAGATGTCGCCTTCTGAT | 250 |
| *1,2DHVR*[a] | 1,2-dihydrovomilenine reductase [N/A] | FW- CGAAGGAGACTCAAGAGATGC<br>RV- CATGCTACTTTAGGACACACAC | 302 |
| *AAE* | Acetylajmalan esterase [AY762990.1] | FW- ATATTCAATGGGACGGTACTC<br>RV- GCGACGGTAAGCAGCTTGTG | 257 |
| *NNMT* | Norajmaline *N*-methyltransferase [KC708449.1] | FW- AGATGGTTGACACAACTTCC<br>RV- GCAAACTCCATAAAGCATTC | 293 |
| *Actin*[b] | Actin [N/A] | FW- GACCTTGAAGTACCCTATTGAG<br>RV- GTACGCCCACTAGCATATAGAG | 249 |
| *60S*[b] | 60S-ribosomal RNA subunit [N/A] | FW- TCTTAGTTGGAATGTTCAGCACCTG<br>RV- CAAGGTTGGAGCCCCTGCTCGTGTT | 198 |

[a] NADPH-dependent reductase cDNA sequences
[b] Reference genes used for normalization of qPCR data.
N/A= Non annotated

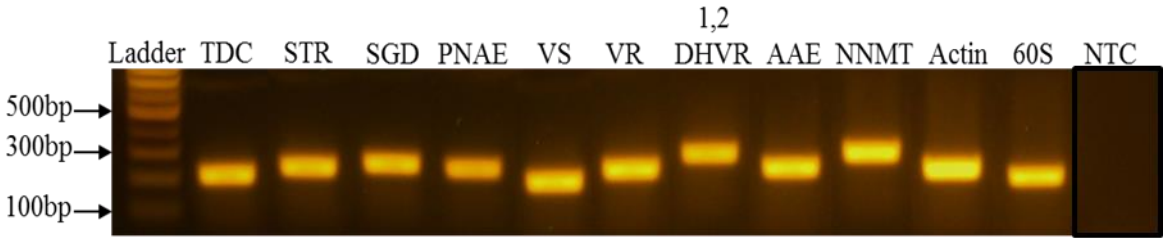



**Figure A-1: Validation of qPCR primers.** Agarose gel electrophoresis displaying DNA fragments produced by PCR using 100 ng of *R. serpentina* roots cDNA as template and the primers listed in **Table A-1**along with no template control (NTC).

### A.2.2 Selection of reference gene for RT-qPCR

An important part of a RT-qPCR analysis is the selection of the appropriate reference gene commonly used for normalization of qPCR data (Radonić et al., 2004). Therefore, we decided to test two of the most common reference genes, actin and 60S ribosomal RNA subunit (**Table A-1**). To assess the stability of each reference gene across multiple organs and leaf developmental stages in *R. serpentina*, we determined the cycle threshold ($C_T$) value in five biological replicates by qPCR (**Figure A-2**). Furthermore, we calculated whether there was a statistically significant difference ($P < 0.05$) between the $C_T$ means across organs and leaf developmental stages analyzed. We found that *actin* transcript levels in terms of $C_T$ values were more homogeneous than *60S* $C_T$ values throughout samples (**Figure A-2**), and the $C_T$ means were not significantly different to each other ($P = 0.1086$). Therefore we decided to use actin as our reference gene for further RT-qPCR experiments.

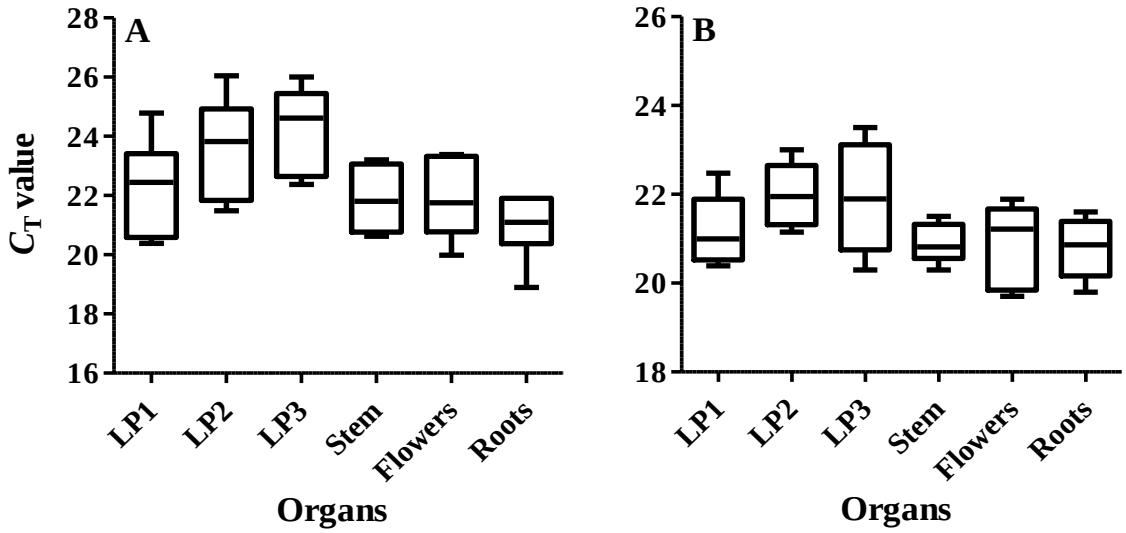



**Figure A-2: Stability of reference genes across organs and leaf developmental stages in *R. serpentina*.** Transcription levels of (**A**) *60S* and (**B**) *actin* in absolute $C_T$ values are shown, boxes indicate 25/75 percentiles, whisker caps indicate maximum and minimum values, the line marks the median. $C_T$ values were obtained from five biological replicates. All qPCR experiments were performed using 10 ng of cDNA. Abbreviations: LP1, youngest leaf pair next to apical meristem; LP2, older leaf pair next to LP1; LP3, older leaf pair next LP2.

### A.2.3 Transcript expression analysis of genes involved in ajmaline biosynthesis

To provide a comprehensive view of the ajmaline metabolic pathway throughout different organs and leaf developmental stages in *R. serpentina*, we determined the transcription levels (in terms of $C_T$ values) of each biosynthetic gene in roots, stem, flowers, and three developmental stages of leaves (**Figure A-3**). Since roots are the main site of ajmaline biosynthesis (**Figure 4-3D**), we decided to calculate the transcript levels of each biosynthetic gene relative to roots and normalize each value to its *actin* $C_T$ counterpart using the following formula for each biological replicate: $\Delta\Delta C_T = 2^{\wedge} - [(C_{T\ \text{biosynthetic gene}} - C_{T\ \text{actin}}) - (C_{T\ \text{biosynthetic gene in roots}} - C_{T\ \text{actin in roots}})]$ (Livak and Schmittgen**,**

2001). Thus, by using this formula we directly compared the biosynthetic gene transcripts in each organ against roots, an organ specialized in producing high levels of ajmaline (**Figure-4-3D**). Our results displayed a decrease of transcript levels in leaves with respect to their age (**Figure A-2**), suggesting that activation of the ajmaline pathway occurs at early stages of leaf development and declines progressively as the leaf matures. We have previously observed this pattern in ajmaline alkaloid accumulation (**Figure-4-3D**). Interestingly, *PNAE* and *AAE* transcripts were highly expressed across all organs and leaf developmental stages (**Figure A-3**).

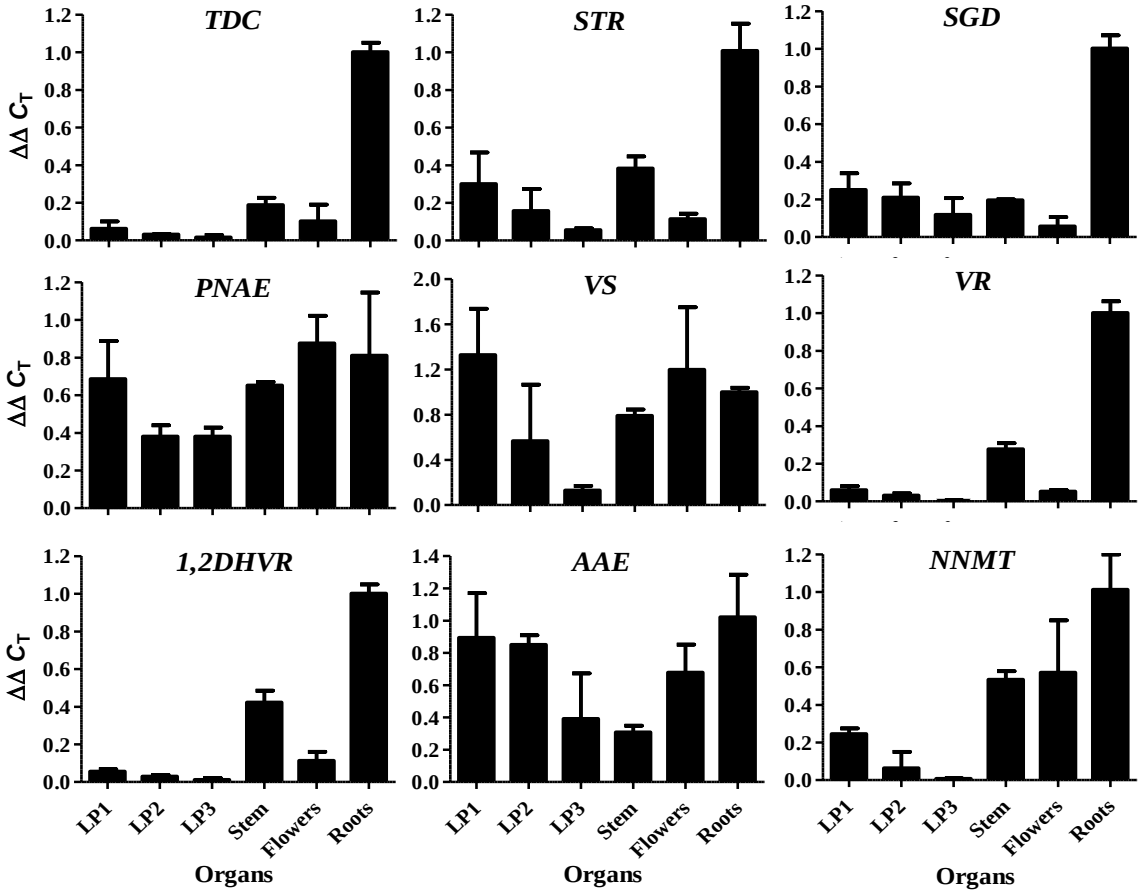



**Figure A-3: Transcript expression analysis.** Relative expression of ajmaline biosynthetic gene transcripts in $\Delta\Delta C_T$ values, bars represent the mean and S.D of three biological replicates. Abbreviations: LP1, youngest leaf pair next to apical meristem;

LP2, older leaf pair next to LP1; LP3, older leaf pair next LP2, Tryptophan decarboxylase (*TDC*), strictosidine synthase (*STR*), strictosidine β-D-glucosidase (*SGD*), polyneuridine aldehyde esterase (*PNAE*), vinorine synthase (*VS*), vomilenine reductase (*VR*), 1,2-dihydrovomilenine reductase (*1,2DHVR*), acetyl ajmalan esterase (*AAE*), norajmaline *N*-methyltransferase (*NNMT*).

In contrast, *TDC*, *VR*, and *1,2DHVR* transcript levels were low in leaves when compared to roots and stem transcripts. Furthermore, *VS* transcripts levels were higher in first leaf pair and flowers compared to roots.

### A.2.4 Alkaloid profile analysis in *R. serpentina* organs

To validate and interpret our transcript expression analysis in terms of alkaloid biosynthesis, we isolated the alkaloid constituents of different organs and leaf developmental stages in *R. serpentina* by performing a methanol extraction followed by acid-base extraction to separate alkaloids from other molecules (see material and methods). Furthermore, we ran these alkaloid extracts through ultra-performance liquid chromatography coupled to a mass detector (UPLC-MS) and identified specific molecules based on their mass and UV absorption spectra using the Dictionary of Alkaloids as a reference (Buckingham et al., 2010). Since ajmaline was the only MIA standard available (Rt = 2.3 min, m/z+327, UV max at 204, 246 and 290 nm), special attention was paid to this MIA for all samples analyzed. As previously described in chapter 4 (**Figure 4-3D**), ajmaline accumulation was highest in roots followed by stem, leaf pair one and flowers (**Figure A-4A**) Interestingly, vomilenine (Rt = 3.1 min, m/z+351, UV max at 221 and 267 nm) was the only MIA involved in the ajmaline

biosynthetic pathway that was identified and it was present in all organs including in leaves of different ages and developmental stages (**Figure A-4B**). Not surprisingly, the reduced forms of vomilenine, 1,2-dihydrovomilenine (m/z+353) and 17-*O*-acetylnorajmaline (m/z+355) were detected in organs where ajmaline biosynthesis is active (**Figure A-5A and B**). These two MIAs (vomilenine and 1,2-dihydrovomilenine) had the same retention times (Rt= 3.1 min). Additionally, roots and stems contained 17-*O*-acetylnorajmaline rather than vomilenine or 1,2-dihydrovomilenine as a prominent MIA (**Figure A-4 and A-5**) and suggested that vomilenine was being converted to 17-*O*-acetylnorajmaline in these organs.

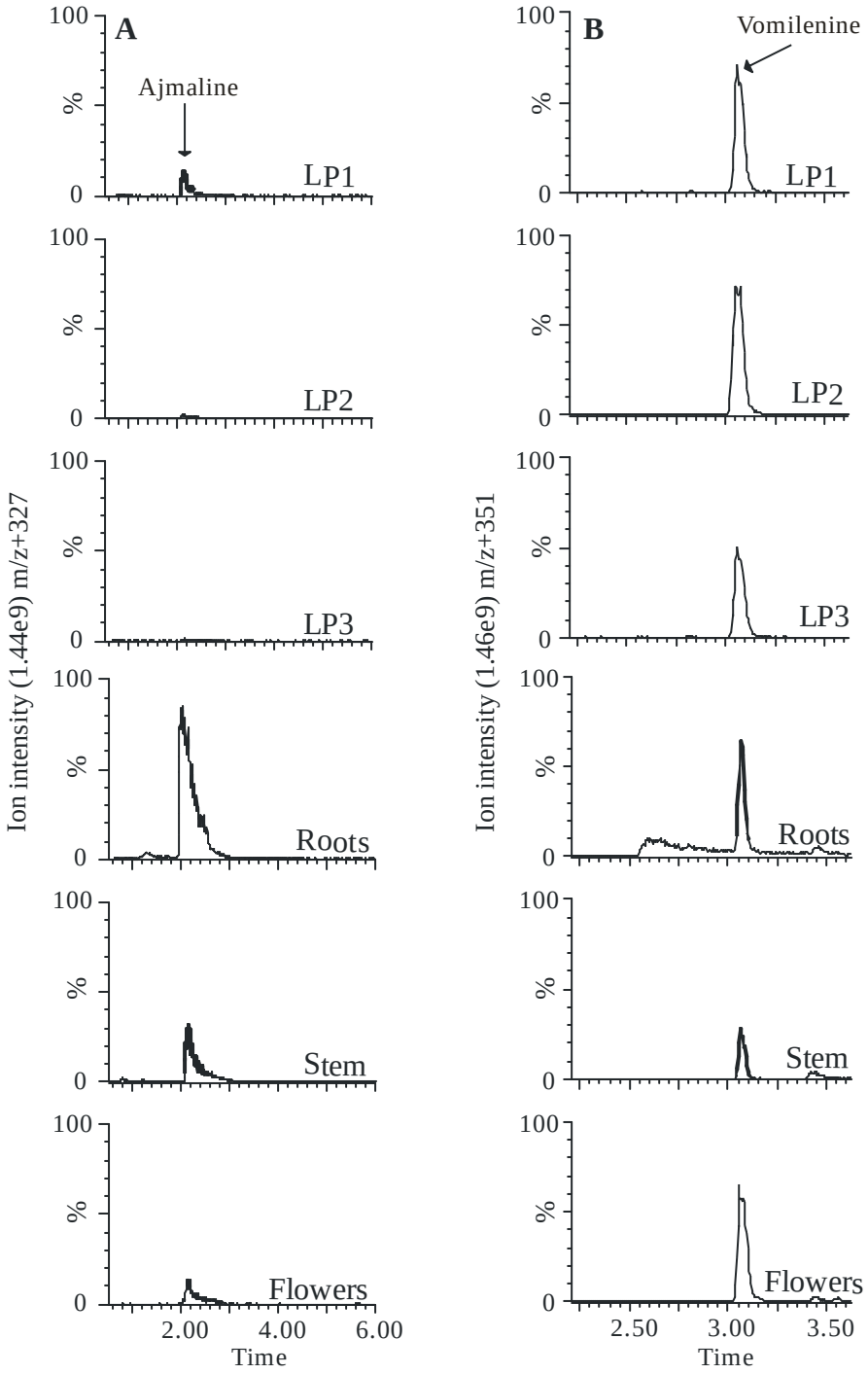



**Figure A-4: Ajmaline and vomilenine alkaloid profile.** UPLC-MS chromatograms displaying the ion intensity of (**A**) ajmaline (Rt = 2.3 min, m/z+327) and (**B**) vomilenine (Rt = 3.1 min, m/z+351) across organs and leaf developmental stages of *R. serpentina*. Selected samples were run in equal concentrations, 20 µg fresh weight per µL.

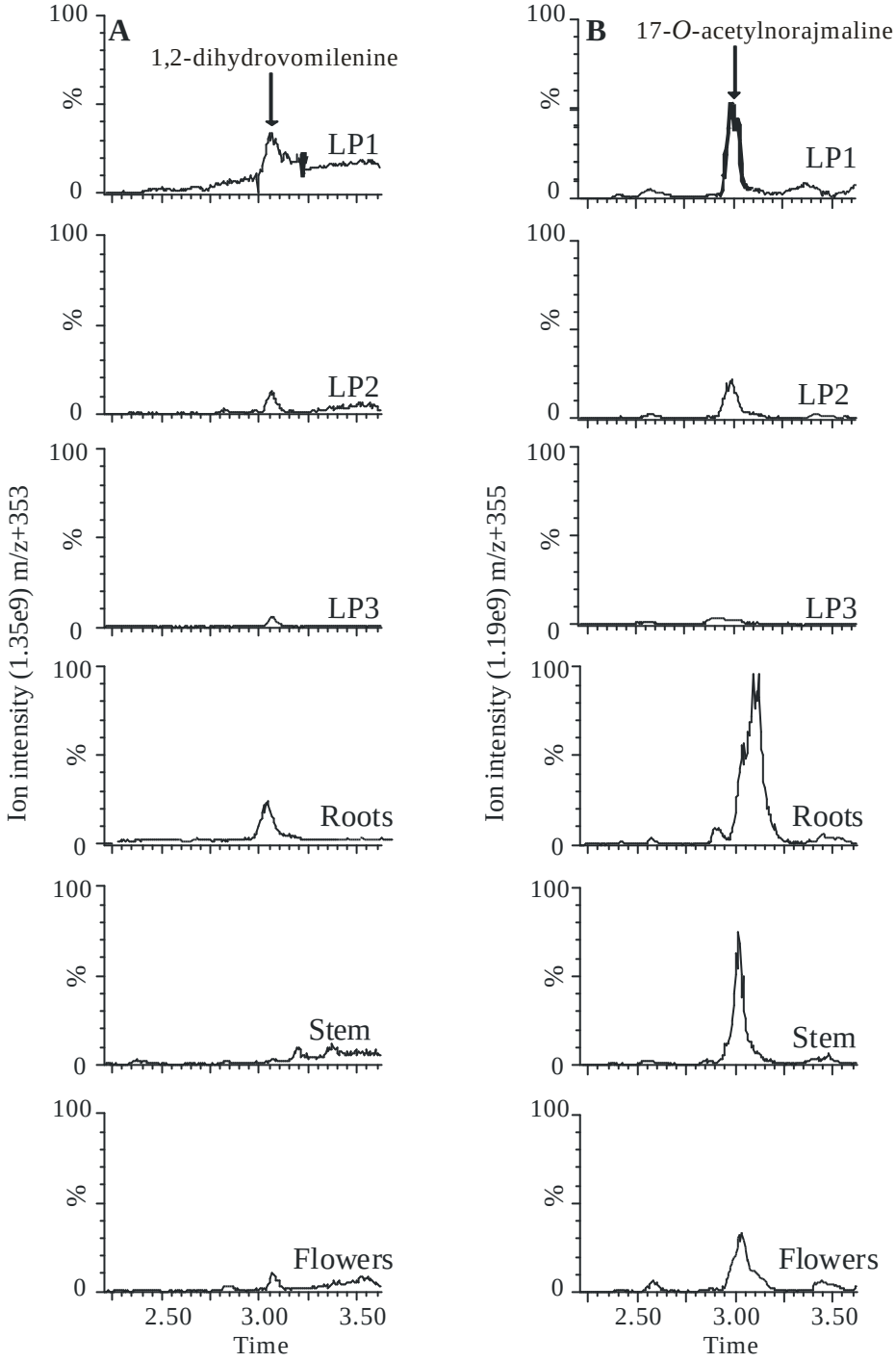



**Figure A-5: 1,2-dihydrovomilenine and 17-*O*-acetylnorajmaline alkaloid profile.** UPLC-MS chromatograms displaying the ion intensity of (**A**) 1,2-dihydrovomilenine (Rt = 3.1 min, m/z+353) and (**B**) 17-*O*-acetylnorajmaline (Rt = 3.1 min, m/z+355) across

organs and leaf developmental stages of *R. serpentina*. Selected samples were run in equal concentrations, 20 µg fresh weight per µL.

### A.3 Discussion

The ajmaline biosynthetic pathway has been extensively studied at the biochemical level in cell suspension cultures of *R. serpentina.* Many of the genes involved in the ajmaline pathway have been cloned and functionally characterized (Ruppert et al., 2005a). The present study uses RT-qPCR analysis and MIA metabolite profiling to correlate ajmaline biosynthesis with the MIAs accumulating in different organs of *R. serpentina* plants. The results obtained reveal that 9 genes in the ajmaline pathway are preferentially expressed in the youngest leaves and that their levels decrease progressively in older leaves (**Figure A-3**). The MIA metabolite profiles of the same tissues suggest that most MIA biosynthesis appears to take place in young leaves and declines within older leaves (**Figure A-4**). The RT-qPCR analysis also indicates that older leaves and flowers may be less active in MIA biosynthesis due to lower transcript levels of the two NADPH-dependent reductases: *VR* and *1,2DHVR* (**Figure A-3**). Although this study provides a general overview of the ajmaline biosynthetic genes transcripts across organs and leaf developmental stages in *R. serpentina*, it is important to bear in mind that metabolite flux is mainly dictated by available pools of intermediates in combination with the concentration of enzymes and their kinetic properties (Fell and Cornish-Bowden, 1997). Additionally, our alkaloid profile analysis of ajmaline intermediates suggests that biosynthesis of vomilenine represents an important branching point towards the production of MIAs with the ajmalan backbone. Moreover, our results suggest that biosynthesis of vomilenine does not appear to be regulated by the particular

developmental stage, which may explain the high transcription levels of genes such as polyneuridine aldehyde esterase (*PNAE*) and vinorine synthase (*VS*), the latter being responsible for the formation of the ajmalan backbone. In conclusion, this study provides the first step towards a broad understanding of the ajmaline biosynthetic pathway in *R. serpentina* plants. Further analysis will require elucidation of the unknown pathway step genes and localization of the metabolic pathway enzymes by *in situ* immunohistochemistry.

## A.4 Materials and methods

### A.4.1 Plant material

*Rauvolfia serpentina* (L.) Benth was grown in artificial soil (Sun Gro Horticulture Canada Ltd.) under controlled greenhouse conditions (22 ± $2^0$C, 70% humidity, 16 h photoperiod) and supplemented once every month with all-purpose fertilizer [Miracle Gro, 12% nitrogen, 4% phosphate, 8% potassium (Scotts Canada)]. Material from 3-month-old plants was used for developmental phytochemical, biochemical and transcriptional analysis (**Figure S4-8**)

### A.4.2 RNA isolation and Reverse-Transcription quantitative-PCR (RT-qPCR)

Total RNA was isolated from different organs of *R. serpentina* using TRIzol® Reagent (Life Technologies) according to the manufacturer's protocol. First strand cDNA synthesis was performed using AMV Reverse Transcriptase (Promega) and oligo(dT) following the manufacturer's protocol. Gene expression experiments were performed using iTaq™ SYBR Green Supermix (Bio-Rad) under the following conditions: 95$^0$C for 3 min, then 40 cycles of 95$^0$C for 15 sec, 55$^0$C for 20 sec, 72$^0$C for 30 sec for each target

gene using approximately 10 ng of cDNA and the qPCR primers listed in **Table A-1**in a final volume of 10 µL. The expression levels were analyzed with the Bio-Rad CFX Manager Software (Bio-Rad), and normalized to the housekeeping gene [*actin* from *R. serpentina*]. To determine relative fold differences, the $C_T$ value for each target gene was normalized to the $C_T$ value for actin, and was calculated relative to a calibrator (roots) using the $2^{-\Delta\Delta C_T}$ method (Livak, K. and Schmittgen, T. 2001). Experiments were conducted in three biological replicates.

### A.4.3 Alkaloid extraction

Plant material from *R. serpentina* (~0.5 g) was ground to a fine powder under liquid $N_2$ and incubated with 5 mL of MeOH at room temperature with wrist shaking (~100 rpm) for 30 min. The extract was centrifuged at 3,000 x *g* for 5 min to pellet cell debris. Subsequently, the supernatant was transferred to a new tube and MeOH was evaporated using the SPD SpeedVac (Thermo Savant). The residue was resuspended in 4 mL of 20% MeOH in water and subjected to acid-base alkaloid extraction performed as described by Murata, J. and De Luca, V., 2005. Finally, 5 µL of each diluted sample was subjected to analysis by UPLC-MS.

### A.4.4 Ultra-Performance Liquid Chromatography analysis

Ultra-performance liquid chromatography analysis was performed using an Acquity Ultra Performance BEH C18 column with a 1.7 µm particle size, 2.1 x 50 mm dimension. The analytes were detected by photodiode array and MS. The solvent system for alkaloid analysis were as follows: solvent A [methanol:acetonitrile:5 mM ammonium acetate (6:14:80 v/v/v)] and solvent B [methanol:acetonitrile:6.2 mM ammonium acetate (25:65:10 v/v/v)]. The following linear elution gradient was used: 0-0.5 min 99% A, 1%

B at 0.3 mL $min^{-1}$; 0.5-0.6 min 99% A, 1% B at 0.4 mL $min^{-1}$; 0.6-7.0 min 1% A, 99% B at 0.4 mL $min^{-1}$; 7.0-8.0 min 1% A, 99% B at 0.4 mL $min^{-1}$; 8.0-8.3 min 99% A, 1% B at 0.4 mL $min^{-1}$; 8.3-8.5 min 99% A, 1% B at 0.3 mL $min^{-1}$; and 8.5-10 min 99% A, 1% B at 0.3 mL $min^{-1}$. The mass spectrometer was operated on positive ion mode with a capillary voltage of 3.10 kV, cone voltage of 48 V, cone gas flow 2 L $h^{-1}$, desolvation gas flow 600 L $h^{-1}$, desolvation temperature of 350 $^0$C, and a source temperature of 150 $^0$C.

**A.5 Acknowledgments**

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (V.D.L.), Canada Research Chairs (V.D.L.), Genome Canada, Genome Alberta, Genome Prairie, Genome British Columbia, the Canada Foundation for Innovation, the Ontario Ministry of Research and Innovation, the National Research Council of Canada and other government and private sector partners, P.C.F. was supported by a postgraduate PhD scholarship from Consejo Nacional de Ciencia y Tecnología (CONACyT), México.

# References


**Achor, R., Hanson, N., and Gifford, R.** (1955). Hypertension treated with *Rauwolfia serpentina* (whole root) and with reserpine: controlled study disclosing occasional severe depression. *Journal of the American Medical Association*, **159**, 841-845.

**Akinloye, B.A., Court, W.E.** (1980a) Leaf alkaloids of *Rauvolfia volkensii*. *Phytochemistry* **19**, 307-311.

**Akinloye, B.A., Court, W.E.** (1980b) Leaf alkaloids of *Rauvolfia oreogiton*. *Phytochemistry* **19**, 2741-2745.

**Arai H., Y. Hirasawa, I. Rahman, N., Zaini, N.C., Sato, S., Aoyama, C., Takeo, J., Morita, H.** (2010) Alstiphyllanines E-H, picraline and ajmaline-type alkaloids from *Alstonia macrophylla* inhibiting sodium glucose transporter. *Bioorg and Med. Chem*, 2152-2158.

**Asada, K., Salim, V., Masada-Atsumi, S., Edmunds, E., Nagatoshi, M., Terasaka, K., Mizukami, H. & De Luca, V.** (2013). A 7-deoxyloganetic acid glucosyltransferase contributes a key step in secologanin biosynthesis in Madagascar periwinkle. *The Plant Cell*, **25**, 4123-4134.

**Baulcombe, D. C.** (1999). Fast forward genetics based on virus-induced gene silencing. *Current opinion in plant biology*, **2**, 109-113.

**Bayer, A., Ma, X., & Stöckigt, J.** (2004). Acetyltransfer in natural product biosynthesis: functional cloning and molecular analysis of vinorine synthase. *Bioorganic & medicinal chemistry*, **12**, 2787-2795.

**Bellet, P., and Van Thuong, T.** (1970). *U.S. Patent No. 3,520,778*. Washington, DC: U.S. Patent and Trademark Office.

**Boga M., Kolak, U., Topcu G., Bahadori F., Kartal M., Farnsworth N.R.** (2011) Two new indole alkaloids from *Vinca herbacea* L. *Phytochemistry Letters* 4, 399-403.

**Bouvier, F., Dogbo, O., & Camara, B.** (2003). Biosynthesis of the food and cosmetic plant pigment bixin (annatto). *Science*, ***300***, 2089-2091.

**Bouvier, F., Suire, C., d'Harlingue, A., Backhaus, R. A., & Camara, B.** (2000). Molecular cloning of geranyl diphosphate synthase and compartmentation of monoterpene synthesis in plant cells. *The Plant Journal*, ***24***, 241-252.

**Bradford, M.** (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72**, 248-254.

**Brugada, J., Brugada, P., and Brugada, R.** (2003). The ajmaline challenge in Brugada syndrome. *European Heart J.*, **24**, 1085-1086.

**Buckingham, J., Baggaley, K., Roberts, A., Szabo, L., eds.** (2010) Dictionary of Alkaloids, CRC Press Taylor & Francis Group, Boca Raton, FL.

**Cavalier-Smith, T.** (1992). Origins of secondary metabolism. *Secondary metabolites: their function and evolution*, pg. 64-87.

**Chatterjee, A., Mukherjee, B., Ray, A.B., Das, B.** (1965) Alkaloid from leaves of *Alstonia scholaris*. *Tetrahedron Lett* **41**, 3633-3637.

**Collu, G., Unver, N., Peltenburg-Looman, A. M., van der Heijden, R., Verpoorte, R., & Memelink, J.** (2001). Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. *FEBS letters*, **508**, 215-220.

**Contin, A., van der Heijden, R., Lefeber, A. W., & Verpoorte, R.** (1998). The iridoid glucoside secologanin is derived from the novel triose phosphate/pyruvate pathway in a *Catharanthus roseus* cell culture. *FEBS letters*, ***434***, 413-416.

**Croteau, R., Kutchan, T. M., & Lewis, N. G.** (2000). Natural products (secondary metabolites). *Biochemistry and molecular biology of plants*, **24**, 1250-1319.

**D'Auria, J. C., Chen, F., & Pichersky, E.** (2003). Chapter eleven The SABATH family of MTs in *Arabidopsis thaliana* and other plant species. *Recent advances in Phytochemistry*, **37**, 253-283.

**De Luca, V., & Cutler, A. J.** (1987). Subcellular localization of enzymes involved in indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiology*, **85**, 1099-1102.

**De Luca, V., & St Pierre, B.** (2000). The cell and developmental biology of alkaloid biosynthesis. *Trends in plant science*, **5**, 168-173.

**De Luca, V., Balsevich, J., Tyler, R. T., Eilert, U., Panchuk, B. D., & Kurz, W. G. W.** (1986). Biosynthesis of indole alkaloids: developmental regulation of the biosynthetic pathway from tabersonine to vindoline in *Catharanthus roseus*. *Journal of Plant Physiology*, **125**, 147-156.

**De Luca, V., Balsevich, J., Tyler, R.T., and Kurz, W.G.W**. (1987). Characterization of a novel *N*-methyltransferase (NMT) from *Catharanthus roseus* plants. *Plant Cell Rep.*, **6**, 458-461.

**De Luca, V., Marineau, C., & Brisson, N.** (1989). Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases. *Proceedings of the National Academy of Sciences*, **86**, 2582-2586.

**De Luca, V., Salim, V., Levac, D., Atsumi, S. M., & Yu, F.** (2012). Discovery and functional analysis of monoterpenoid indole alkaloid pathways in plants. *Methods in Enzymology*, **515**, 207-229.

**De Luca, V., Salim, V., Levac, D., Atsumi, S.M., and Yu, F.** (2012). Discovery and functional analysis of monoterpenoid indole alkaloid pathways in plants. *Methods Enzymol.*,**515**, 207-229.

**Dethier, M., De Luca, V.** (1993) Partial purification of an *N*-methyltransferase involved in vindoline biosynthesis in *Catharanthus roseus*. *Phytochemistry*, **32**, 673-678.

**Dixon, R. A., & Paiva, N. L.** (1995). Stress-induced phenylpropanoid metabolism. *The plant cell*, **7**, 1085.

**Dogru, E., Warzecha, H., Seibel, F., Haebel, S., Lottspeich, F., & Stöckigt, J.** (2000). The gene encoding polyneuridine aldehyde esterase of monoterpenoid indole alkaloid biosynthesis in plants is an ortholog of the α/β hydrolase super family. *European Journal of Biochemistry*, **267**, 1397-1406.

**Edgar, R.C.** (2004), MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Research*, **32**, 1792-1797.

**Exkermann, R., Gaich, T.** (2013) The akuammiline alkaloids; origin and synthesis. *Synthesis* **45**, 2813-2823.

**Facchini, P. J.** (2001). Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Annual review of plant biology*, **52**, 29-66.

**Facchini, P. J., & De Luca, V**. (2008). Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. *The Plant Journal*, **54**, 763-784.

**Facchini, P. J., Bohlmann, J., Covello, P. S., De Luca, V., Mahadevan, R., Page, J. E., Ro, D.K., Sensen, C. W., Storms, R., Martin, V.J.** (2012) Synthetic biosystems for the production of high-value plant metabolites. *Trends in Biotech*, **30**, 127-131.

**Falkenhagen, H., & Stöckigt, J.** (1995). Enzymatic Biosynthesis of Vomilenine, a Key Intermediate of the Ajmaline Pathway, Catalyzed by a Novel Cytochrome P 450-Dependent Enzyme from Plant Cell Cultures of *Rauwolfia serpentina*. *Zeitschrift für Naturforschung C*, **50**, 45-53.

**Fell, D., & Cornish-Bowden, A.** (1997). *Understanding the control of metabolism* **Vol. 2**. London: Portland press.

**Felsenstein, J.** (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**, 783-791.

**Gao, S., von Schumann, G., & Stöckigt, J.** (2002). A newly-detected reductase from Rauvolfia closes a gap in the biosynthesis of the antiarrhythmic alkaloid ajmaline. *Planta medica*, **68**, 906-911.

**Geu-Flores, F., Sherden, N.H., Courdavault, V., Burlat, V., Glenn, W.S., Wu, C., Nims, E., Cui, Y. & O'Connor, S.E.** (2012). An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. *Nature*,**492**,138-142.

**Góngora-Castillo, E., Fedewa, G., Yeo, Y., Chappell, J., DellaPenna, D., & Buell, C. R.** (2012). Genomic approaches for interrogating the biochemistry of medicinal plant species. *Methods in enzymology*, **517**, 139.

**Grossmann, E., Sevcovic, P., Szasz, K.** (1973) Picrinine in *Vinca minor*. *Phytochemistry* **12**, 2058.

**Hai C.X., Ya-Ping L., Tao F., Xiao-Dong L.** (2008) Picrinine-type alkaloids from the leaves of *Alstonia scholaris*. *J. Chinese J of Nat. Med*. **6**, 20-22.

**Hartmann, T.** (2004). Plant-derived secondary metabolites as defensive chemicals in herbivorous insects: a case study in chemical ecology. *Planta*, **219**, 1-4.

**Hartmann, T.** (2007). From waste products to ecochemicals: fifty years research of plant secondary metabolism. *Phytochemistry*, **68**, 2831-2846.

**Harvey, A. L., Edrada-Ebel, R., & Quinn, R. J.** (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nature Reviews Drug Discovery*, **14**, 111-129.

**Herbert, R. B.** (1989). *The biosynthesis of secondary metabolites*. Springer Science & Business Media.

**Hibi, N., Fujita, T., Hatano, M., Hashimoto, T., & Yamada, Y.** (1992). Putrescine N-Methyltransferase in Cultured Roots of *Hyoscyamus albus* n-Butylamine as a Potent Inhibitor of the Transferase both in Vitro and in Vivo. *Plant Physiology*, **100**, 826-835.

**Horton, P., Park, K., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K.** (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Research*, **259**, 585-587.

**Iijima, Y., Gang, D. R., Fridman, E., Lewinsohn, E., & Pichersky, E.** (2004). Characterization of geraniol synthase from the peltate glands of sweet basil. *Plant Physiology*, **134**, 370-379.

**Ikeda, H., Esaki, N., Nakai, S., Hashimoto, K., Uesato, S., Soda, K., & Fujita, T.** (1991). Acyclic monoterpene primary alcohol: NADP+ oxidoreductase of *Rauwolfia serpentina* cells: the key enzyme in biosynthesis of monoterpene alcohols. *Journal of biochemistry*, **109**, 341-347.

**Ingham, J. L., Koskinen, A., & Lounasmaa, M.** (2012). *Fortschritte der Chemie organischer Naturstoffe/Progress in the Chemistry of Organic Natural Products*, **43**. Springer Science & Business Media.

**Irmler, S., Schröder, G., St-Pierre, B., Crouch, N.P., Hotze, M., Schmidt, J., Strack, D., Matern, U. & Schröder, J.,** (2000). Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. *The Plant Journal*, **24**,797-804.

**Itoh, A., Kumashiro, T., Yamaguchi, M., Nagakura, N., Mizushina, Y., Nishi, T., and Tanahashi, T.** (2005). Indole alkaloids and other constituents of *Rauwolfia serpentina*. *J. Nat. Prod.*, **68**, 848-852.

**Joshi, C. P., & Chiang, V. L.** (1998). Conserved sequence motifs in plant S-adenosyl-L-methionine-dependent methyltransferases. *Plant molecular biology*, ***37***, 663-674.

**Joyard, J., Ferro, M., Masselon, C., Seigneurin-Berny, D., Salvia, D., Garin, J., Rolland, N.** (2009) Chloroplast proteomics and the compartmentation of plastidial isoprenoid biosynthetic pathways. *Molecular plant*, **2**, 1154-1180.

**Kisakurek, M. V., Leeuwenberg, A. J., & Hesse, M**. (1983). A chemotaxonomic investigation of the plant families of Apocynaceae, Loganiaceae, and Rubiaceae by their indole alkaloid content. *Pelletier, S, W ed (s). Alkaloids: chemical and biological perspectives*, **1**, 211-376.

**Klyushnichenko, V. E., Yakimov, S. A., Tuzova, T. P., Syagailo, Y. V., Kuzovkina, I. N., Wulfson, A. N., & Miroshnikov, A. I.** (1995). Determination of indole alkaloids from *R. serpentina* and *R. vomitoria* by high-performance liquid chromatography and high-performance thin-layer chromatography. *Journal of Chromatography*, **704**, 357-362.

**Kozbial, P. Z., & Mushegian, A. R.** (2005). Natural history of S-adenosylmethionine-binding proteins. *BMC structural biology*, **5**, 1-26.

**Kutchan, T. M., Hampp, N., Lottspeich, F., Beyreuther, K., & Zenk, M. H.** (1988). The cDNA clone for strictosidine synthase from *Rauvolfia serpentina* DNA sequence determination and expression in Escherichia coli. *FEBS letters*, **237**, 40-44.

**Levac, D., Cázares, P., Yu, F., and De Luca V.** (2016) A picrinine-*N*-methyltransferase belongs to a new family of γ-tocopherol-like *N*-methyltransferases found in medicinal plants that make biologically active monoterpenoid indole alkaloids. *Plant Physiology* (in press).

**Levac, D., Murata, J., Kim, W., De Luca, V.** (2008) Application of carborundum abrasion for investigation of leaf epidermis: molecular cloning of *Catharanthus roseus* 16-hydroxytabersonsine-16-O-methyltransferase, *The Plant Journal*, **53**, 225-236.

**Levac, Dylan** (2013). "The γ-tocopherol-like family of *N*-methyltransferases: A taxonomically clustered gene family encoding enzymes responsible for *N*-methylation of monoterpene indole alkaloids" *PhD Thesis*, Brock University.

**Liscombe, D. K., & O'Connor, S. E.** (2011). A virus-induced gene silencing approach to understanding alkaloid metabolism in *Catharanthus roseus*. *Phytochemistry*, **72**, 1969-1977.

**Liscombe, D. K., Louie, G. V., & Noel, J. P.** (2012). Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Natural product reports*, **29**, 1238-1250.

**Liscombe, D. K., Usera, A. R., & O'Connor, S. E.** (2010). Homolog of tocopherol C methyltransferases catalyzes N methylation in anticancer alkaloid biosynthesis. *Proceedings of the National Academy of Sciences*, **107**, 18793-18798.

**Liscombe, D.K., and Facchini, P.J.** (2007). Molecular cloning and characterization of tetrahydroprotoberberine cis-*N*-methyltransferase, an enzyme involved in alkaloid biosynthesis in opium poppy. *J. Biol. Chem.*, **282**, 14741-14751.

**Liu, D. L., & Lovett, J. V.** (1993). Biologically active secondary metabolites of barley. II. Phytotoxicity of barley allelochemicals. *Journal of Chemical Ecology*, ***19***, 2231-2244.

**Liu, H., Wu, B., Zheng, Q., Feng, X.** (1991) New indole alkaloids from *Amsonia sinensis*. *Planta Medica*, **57**, 566-568.

**Livak, K.J., and Schmittgen, T.D.** (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*, **25**, 402-408.

**Luijendijk, T. J., Stevens, L. H., & Verpoorte, R.** (1998). Purification and characterisation of strictosidine β-D-glucosidase from *Catharanthus roseus* cell suspension cultures. *Plant Physiology and Biochemistry*, **36**, 419-425

**Ma, X., Koepke, J., Panjikar, S., Fritzsch, G., & Stöckigt, J.** (2005). Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *Journal of Biological Chemistry*, **280**, 13576-13583.

**Maresh, J.J., Giddings, L.A., Friedrich, A., Loris, E.A., Panjikar, S., Trout, B.L., Stöckigt, J., Peters, B. & O'Connor, S.E.** (2008). Strictosidine synthase: mechanism of a Pictet-Spengler catalyzing enzyme. *Journal of the American Chemical Society*, **130**, 710-723.

**Mattern-Dogru, E., Ma, X., Hartmann, J., Decker, H., & Stöckigt, J.** (2002). Potential active-site residues in polyneuridine aldehyde esterase, a central enzyme of indole alkaloid biosynthesis, by modelling and site-directed mutagenesis. *European Journal of Biochemistry*, **269**, 2889-2896.

**Meister, G., & Tuschl, T.** (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature*, **431**, 343-349.

**Murata, J., and Luca, V. D.** (2005). Localization of tabersonine 16-hydroxylase and 16-OH tabersonine-16-O-methyltransferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in *Catharanthus roseus*. *Plant J.*, **44**, 581-594.

**Murata, J., Bienzle, D., Brandle, J.E., Sensen, C.W. and De Luca, V.** (2006) Expressed sequence tags from Madagascar periwinkle (*Catharanthus roseus*). *FEBS Letters*, **580**, 4501-4507.

**Murata, J., Roepke, J., Gordon, H., & De Luca, V.** (2008). The leaf epidermome of *Catharanthus roseus* reveals its biochemical specialization. *The Plant Cell*, **20**, 524-542.

**Nagakura, N., Rüffer, M., & Zenk, M. H**. (1979). The biosynthesis of monoterpenoid indole alkaloids from strictosidine. *Journal of the Chemical Society, Perkin Transactions* **1**, 2308-2312.

**Nei, M., Kumar, S.** (2000) Molecular Evolution and Phylogenetics. Oxford University Press, New York.

**Néron, B., Ménager, H., Maufrais, C., Joly, N., Maupetit, J., Letort, S., Carrere, S., Tuffery, P. and Letondal, C.** (2009). Mobyle: a new full web bioinformatics framework. *Bioinformatics*, **25**, 3005-3011.

**Noble, R. L.** (1990). The discovery of the vinca alkaloids-chemotherapeutic agents against cancer. *Biochemistry and cell biology*, **68**, 1344-1351.

**O'Connor, S. E., & Maresh, J. J.** (2006). Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Natural product reports*, **23**, 532-547.

**O'Hagan, D., & Schmidberger, J. W.** (2010). Enzymes that catalyse SN2 reaction mechanisms. *Natural product reports*, **27**, 900-918.

**Pfitzner, A., & Stöckigt, J.** (1983). Characterization of polyneuridine aldehyde esterase, a key enzyme in the biosynthesis of sarpagine/ajmaline type alkaloids. *Planta medica*, **48**, 221-227.

**Pfitzner, A., Polz, L., & Stöckigt, J.** (1986). Properties of vinorine synthase—the Rauwolfia enzyme involved in the formation of the ajmaline skeleton .*Zeitschrift für Naturforschung C*, **41**, 103-114.

**Pichersky, E., & Gang, D. R.** (2000). Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends in plant science*, **5**, 439-445.

**Pichersky, E., & Gershenzon, J.** (2002). The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Current opinion in plant biology*, ***5***, 237-243.

**Pinçon, G., Maury, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B., & Legrand, M.** (2001). Repression of O-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth. *Phytochemistry*, **57**, 1167-1176.

**Polz, L., Schübel, H., & Stöckigt, J.** (1987). Characterization of 2β (R)-17-O-AcetylajmaIan: Acetylesterase—a Specific Enzyme Involved in the Biosynthesis of the *Rauwolfia Alkaloid* Ajmaline. *Zeitschrift für Naturforschung C*, **42**, 333-342.

**Rakhimov, D.A., Malikov, V.M., Yunusov, S.Y.** (1967) Isolation of Kopsinilam and ervincine Khimiya Prirodnykh Soedinenii, **3**, 354-355.

**Roepke, J., Salim, V., Wu, M., Thamm, A., Murata, J., Ploss, K., Boland, W., De Luca, V.** (2010) Vinca drug components accumulate exclusively in leaf exudates of Madagascar perisinkle. *PNAS*, **107**, 15287-15292.

**Roja, P.C., Sipahimalani, A.T., Heble, M.R., and Chadha, M.S.** (1987). Multiple shoot cultures of *Rauvolfia serpentina*: growth and alkaloid production. *J. Nat. Prod*, **50**, 872-875.

**Ruppert, M., Ma, X., & Stöckigt, J.** (2005a). Alkaloid biosynthesis in Rauvolfia: cDNA cloning of major enzymes of the ajmaline pathway. *Current Organic Chemistry*, **9**, 1431-1444.

**Ruppert, M., Woll, J., Giritch, A., Genady, E., Ma, X., & Stöckigt, J.** (2005b). Functional expression of an ajmaline pathway-specific esterase from Rauvolfia in a novel plant-virus expression system. *Planta*, **222**, 888-898.

**Salim, V., & De Luca, V.** (2013). Towards complete elucidation of monoterpene indole alkaloid biosynthesis pathway: *Catharanthus roseus* as a pioneer system. *Advances in Botanical Research*, **68**, 1-37.

**Salim, V., De Luca, V.** (2013) Towards complete elucidation of monoterpenoid indole alkaloid biosynthesis pathway: *Catharanthus roseus* as a pioneer system. *Annals of Botanical Research*, **68**, 1-38.

**Salim, V., Wiens, B., Masada-Atsumi, S., Yu, F., & De Luca, V.** (2014). 7-deoxyloganetic acid synthase catalyzes a key 3 step oxidation to form 7-deoxyloganetic acid in *Catharanthus roseus* iridoid biosynthesis. *Phytochemistry*, **101**, 23-31.

**Salim, V., Yu, F., Altarejos, J., & Luca, V.** (2013). Virus-induced gene silencing identifies *Catharanthus roseus* 7-deoxyloganic acid-7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis. *The Plant Journal*, **76**, 754-765.

**Sambrock, J., Fritsch, E.F., and Maniatis, T.** (1989). Molecular Cloning: A Laboratory Manual Cold Spring Harbor Laboratory.

**Schmidt, D., & Stöckigt, J.** (1995). Enzymatic formation of the sarpagan-bridge: a key step in the biosynthesis of sarpagine-and ajmaline-type alkaloids. *Planta medica*, **61**, 254-258.

**Seo, H. S., Song, J. T., Cheong, J. J., Lee, Y. H., Lee, Y. W., Hwang, I. & Do Choi, Y.** (2001). Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses. *Proceedings of the National Academy of Sciences*, **98**, 4788-4793.

**Shangb, J., Caia, X., Fenga, T., Zhaob, Y., Wangb, J., Zhangc, L., Yanc, M., Luoa, D.** (2010) Pharmacological evaluation of *Alstonia scholaris*: anti-inflammatory and analgesic effects, *Journal of Ethnopharmacology*, **129**, 174-181.

**Sneath, P.H.A., Sokal, R.R.** (1973) Numerical Taxonomy. Freeman, San Francisco.

**Stöckigt, D., Unger, M., Belder, D., & Stöckigt, J.** (1997). Analysis of *Rauwolfia alkaloids* employing capillary electrophoresis-mass spectrometry.*Natural Product Letters*, **9**, 265-272.

**Stöckigt, J., & Panjikar, S.** (2007). Structural biology in plant natural product biosynthesis—architecture of enzymes from monoterpenoid indole and tropane alkaloid biosynthesis. *Natural product reports*, **24**, 1382-1400.

**Stöckigt, J., Pfitzner, A., & Firl, J.** (1981). Indole alkaloids from cell suspension cultures of *Rauwolfia serpentina* Benth. *Plant Cell Reports*, **1**, 36-39.

**Stöckigt, J., Pfitzner, A., & Keller, P. J.** (1983). Enzymatic formation of ajmaline. *Tetrahedron Letters*, **24**, 2485-2486.

**St-Pierre, B., Vazquez-Flota, F. A., & De Luca, V.** (1999). Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate. *The Plant Cell*, **11**, 887-900.

**Szabó, L.F.** (2008) Ridorous biogenetic network for a group of indole alkaloids derived from strictosidine. *Molecules*, **13**, 1875-1896.

**Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S.** (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28**, 2731-2739.

**Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S.** (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, **30**, 2725-2729.

**Theis, N., & Lerdau, M.** (2003). The evolution of function in plant secondary metabolites. *International Journal of Plant Sciences*, **164**, S93-S102.

**Thompson, J.D., Higgins, D.G. and Gibson, T.J.** (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673-4680

**Tietze, L. F.** (1983). Secologanin, a biogenetic key compound—synthesis and biogenesis of the iridoid and secoiridoid glycosides. *Angewandte Chemie International Edition in English*, **22**, 828-841.

**Treimer, J. F., & Zenk, M. H.** (1979). Purification and properties of strictosidine synthase, the key enzyme in indole alkaloid formation. *European Journal of Biochemistry*, **101**, 225-233.

**Uefuji, H., Ogita, S., Yamaguchi, Y., Koizumi, N., & Sano, H.** (2003). Molecular cloning and functional characterization of three distinct N-methyltransferases involved in the caffeine biosynthetic pathway in coffee plants. *Plant Physiology*, **132**, 372-380.

**Verpoorte, R., & Alfermann, A. W.** (2000). *Metabolic engineering of plant secondary metabolism*. Springer Science & Business Media.

**Vetter, H.-P., Mangold, U., Schröder, G., Marner, F.-J., Werck- Reichhart, D. & Schröder, J.** (1992) Molecular analysis and heterologous expression of an inducible cytochrome P-450 protein from periwinkle (*Catharanthus roseus* L.). *Plant Physiology*, **100**, 998-1007.

**Vogt, T.** (2010). Phenylpropanoid biosynthesis. *Molecular plant*, **3**, 2-20.

**von Schumann, G., Gao, S. and Stöckigt, J.** (2002). Vomilenine reductase—a novel enzyme catalyzing a crucial step in the biosynthesis of the therapeutically applied antiarrhythmic alkaloid ajmaline. *Bioorganic & medicinal chemistry*, **10**, 1913-1918.

**Vranová, E., Coman, D., & Gruissem, W.** (2013). Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annual review of plant biology*, ***64***, 665-700.

**Wink, M.** (2003). Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry*, ***64***, 3-19.

**Wu F., Kerčmar P, Zhang C. and Stöckigt, J.** (2016) Sarpagan-Ajmalan-Type Indoles: Biosynthesis, Structural Biology, and Chemo-Enzymatic Significance. *Alkaloids Chem. Biol.* **76**: 1-61.

**Xiao M., Zhang Y., Chen X., Lee E.J., Barber C., Chakrabarty R., Desgagné-Penix I., Haslam T., Kim Y.-B., Liu E., MacNervin G., Masada-Atsumi S., Reed D., Stout J., Zerbe P., Zhang Y., Bohlmann J., Covello P., De Luca V., Page J., Ro D.-K., Martin V., Facchini P., and Sensen C.W.** (2013). Transcriptome

analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *J. Biotech.*, **166**, 122-134.

**Yang, L., Hill, M., Wang, M., Panjikar, S., & Stöckigt, J**. (2009). Structural Basis and Enzymatic Mechanism of the Biosynthesis of C9-from C10-Monoterpenoid Indole Alkaloids. *Angewandte Chemie International Edition*, **48,** 5211-5213.

**Zuckerkandl, E., Pauling, L.** (1965) Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H.J. Vogel, pp. 97-166. Academic Press, New York.